

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY**

**A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

**ÚSTAV TELEKOMUNIKACÍ**

DEPARTMENT OF TELECOMMUNICATIONS

**MIKROFONNÍ POLE PRO ODHAD SMĚRU  
PŘÍCHÁZEJÍCÍHO ZVUKU**

MICROPHONE ARRAY FOR ESTIMATION OF DIRECTION OF ARRIVAL OF SOUND

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

Ladislav Kubišta

**VEDOUCÍ PRÁCE**

SUPERVISOR

doc. Ing. Jiří Schimmel, Ph.D.

**BRNO 2017**



# Bakalářská práce

bakalářský studijní obor **Audio inženýrství**

Ústav telekomunikací

**Student:** Ladislav Kubišta

**ID:** 174460

**Ročník:** 3

**Akademický rok:** 2016/17

## NÁZEV TÉMATU:

### Mikrofonní pole pro odhad směru přicházejícího zvuku

## POKYNY PRO VYPRACOVÁNÍ:

Prostudujte metody odhadu směru přicházejícího zvuku pomocí mikrofonních polí malých rozměrů založené na měření fázového posunu, korelace a energie a vybrané metody realizujte v prostředí Matlab. Navrhněte a zkonstruuje mikrofonní pole malých rozměrů a proveďte srovnávací měření ve volném poli a v reálném prostředí pro vybrané typy zvukových signálů, např. výstřel, rozbití skla a siréna. Určete přesnost a spolehlivost jednotlivých metod.

## DOPORUČENÁ LITERATURA:

[1] Kinsler, L., Frey, A., Coppens, A., Sanders, J. Fundamentals of Acoustics, JohnWiley and Sons, New York, NY, USA, 1982.

[2] Dan Li., Yu Hen Hu, "Energy-Based Collaborative Source Localization Using Acoustic Microsensor Array," EURASIP Journal on Applied Signal Processing, vol. 2003, no. 4, pp. 321-337, 2003.

[3] Svaizer, P., Matassoni, M., Omologo, M. "Acoustic Source Location in a Three-Dimensional Space Using Crosspower Spectrum Phase," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'97) - Volume 1, 1997.

**Termín zadání:** 1.2.2017

**Termín odevzdání:** 8.6.2017

**Vedoucí práce:** doc. Ing. Jiří Schimmel, Ph.D.

**Konzultant:**

**doc. Ing. Jiří Mišurec, CSc.**  
*předseda oborové rady*

## UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## ABSTRAKT

Tato bakalářská práce se zabývá detekcí směru přicházejícího zvuku pomocí analýzy zvuku. Využívá především metod založených na časovém zpoždění signálu. Algoritmus nalezení směru vychází z křížové korelace a z jejích metod. V závěru jsou shrnuty výsledky měření určování směru, jak naprogramovaného zvuku, tak i zvuku nahraného laboratorních podmínkách a reálném prostředí. Veškeré výpočty byly provedeny pomocí simulačního prostředí Matlab.

## KLÍČOVÁ SLOVA

Detekce směru, zvuk, korelace, mikrofon, Matlab, časové zpoždění.

## ABSTRACT

This bachelor's work describe detection of direction receiving sound by analizing of the sound. Work is based on methods of time delay estimation. Programmed algorithm of estimation direction works on cross-correlation and some selected cross-correlation methods. Results of measuring, as programming sound, so sound recorded in laboratory conditions and real enviroment, are mentioned in conclusion. All calculations were done by platform Matlab

## KEYWORDS

Detection of direction, sound, correlation, microphone, Matlab, time delay.

KUBIŠTA, Ladislav *Mikrofonní pole pro detekci směru přicházejícího zvuku*: bakalářská práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2016. 53 s. Vedoucí práce byl Ing. Jiří Schimmel, Ph.D.



## PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Mikrofonní pole pro detekci směru přicházejícího zvuku“ jsem vypracoval(a) samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno .....

.....

podpis autora(-ky)

## PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Jiřímu Schimmelovi, Ph.D. za trpělivost, pomoc při řešení složitějších problémů, konzultace a podnětné návrhy k práci.

Brno .....

.....

podpis autora(-ky)

# OBSAH

Úvod	10
<b>1 Teoretická část práce</b>	<b>11</b>
1.1 Zvuk . . . . .	11
1.1.1 Šíření zvuku prostorem . . . . .	11
1.1.2 Akustický tlak . . . . .	11
1.1.3 Akustická impedance . . . . .	11
1.1.4 Typy zvukových polí . . . . .	12
1.2 Lokalizace směru přicházejícího zvuku . . . . .	13
1.2.1 Metody pro určování polohy zvuku pomocí časového zpoždění . . . . .	13
1.2.2 Určování polohy zvuku pomocí analýzy energického spektra . . . . .	18
1.3 Matematické funkce . . . . .	21
1.3.1 Diskrétní Fourierova transformace (DFT) . . . . .	21
1.3.2 Rychlá Fourierova transformace (FFT) . . . . .	22
1.3.3 Krátkodobá Fourierova transformace (STFT) . . . . .	22
1.3.4 Křížová korelace . . . . .	24
1.4 Matlab . . . . .	25
<b>2 Vyhodnocení praktické části</b>	<b>26</b>
2.1 Popis měření a zpracování výsledků . . . . .	26
2.1.1 Parametry měření bez vyrobeného mikrofonního pole . . . . .	26
2.1.2 Parametry měření s vyrobeným mikrofonním polem . . . . .	27
2.1.3 Zpracování v Matlabu . . . . .	29
2.2 Výsledky . . . . .	31
2.2.1 Výsledky zkušebního signálu . . . . .	31
2.2.2 Výsledky s reálným signálem bez vyrobeného pole . . . . .	33
2.2.3 Výsledky s reálným signálem a s vyrobeným polem . . . . .	37
<b>3 Závěr</b>	<b>45</b>
Literatura	46
Seznam příloh	48
A Zdrojové výpisy z Matlabu	49
B Obsah přiloženého CD	52
Seznam symbolů, veličin a zkratk	53

# SEZNAM OBRÁZKŮ

1.1	Ukázka nahraného zvuku . . . . .	12
1.2	Časové zpoždění mezi dvěma mikrofony . . . . .	13
1.3	Zvuk dopadající na dva mikrofony . . . . .	15
1.4	Ukázka nastavení mikrofونů pro B-formát. . . . .	19
1.5	Schéma GCC korelace . . . . .	25
2.1	Navržené mikrofonní pole . . . . .	27
2.2	Schéma rozestavení mikrofونů v křížovém poli. . . . .	28
2.3	Schéma rozestavení mikrofونů v kruhovém poli. . . . .	29
2.4	Graf korelovaného naprogramovaného zvuku . . . . .	32
2.5	Korelace zašuměného signálu . . . . .	35
2.6	Graf obecné křížové korelace . . . . .	35
2.7	Graf korelace metodou PHAT . . . . .	36
2.8	Graf korelace metodou ML . . . . .	36
2.9	Grafy průměrné odchylky, měření v laboratoři . . . . .	40
2.10	Grafy průměrné odchylky, měření v laboratoři . . . . .	41
2.11	Grafy průměrné odchylky, měření v laboratoři . . . . .	42
2.12	Grafy průměrné odchylky, měření v reálném prostředí . . . . .	44

## SEZNAM TABULEK

2.1	Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 48 kHz.	33
2.2	Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 96 kHz.	34
2.3	Tabulka výsledků měření v laboratoři, přehled jednotlivých metod . .	39
2.4	Výsledky měření v reálném prostředí, přehled jednotlivých metod. . .	43
2.5	Přehled jednotlivých mikrofonů v poli. . . . .	43

## SEZNAM VÝPISŮ

A.1	GCC metoda naprogramovaná v Matlabu. . . . .	49
A.2	Část kódu naprogramované metody PHAT v Matlabu. . . . .	50
A.3	Část kódu naprogramované metody ML v Matlabu. . . . .	50
A.4	Jednoduché naprogramování umělého zvuku v Matlabu. . . . .	50
A.5	Energetická metoda naprogramovaná v Matlabu . . . . .	51

# ÚVOD

Práce uvádí do problematiky detekce pohybu pomocí analýzy zvuku. V první části jsou rozebrány teoretické předpoklady a postupy pro měření. Jako první se zaměřuji na vlastnosti zvuku, které mohou ovlivnit výsledek mé práce. Dále již přistupuji k návrhu odhadu směru příchozího zvuku pomocí časového zpoždění mezi mikrofony. Jsou zde použity matematické algoritmy jako korelace, rychlá Fourierova transformace (FFT), které dále více popisuji v samostatné kapitole. Popisuji také různé typy mikrofonních polí. V poslední kapitole najdeme výsledky, například naprogramování funkce pro detekci směru přicházejícího zvuku. Je zde také zamyšlení nad volbou jednotlivých proměnných jako vzdálenost mikrofونů od sebe, či volbou správného vzorkovacího kmitočtu. Dále jsou zde výsledné tabulky měření, jak z bezodrazové komory, tak i z reálného prostředí.

# 1 TEORETICKÁ ČÁST PRÁCE

## 1.1 Zvuk

### 1.1.1 Šíření zvuku prostorem

Zvuk je v podstatě kmitání molekul v prostoru, nebo-li kmitání hmotného prostředí. Vlastnosti jeho šíření závisí na jednotlivých vazbách molekul v prostředí (ve vzduchu se zvuk šíří pomaleji, protože jsou molekuly vzduchu dál od sebe než například ve vodě), tudíž třeba ve vakuu nemůže zvuk ani existovat. Když je prostor neomezený a bezodrazový, označujeme jej za volný.

#### Rychlost šíření zvuku ve vzduchu

Za předpokladu, že suchý vzduch má barometrický tlak, tj.  $p_{00} = 1013 \text{ kPa}$ , potom rychlost zvuku ve vzduchu o teplotě  $t$  (ve  $^{\circ}\text{C}$ ) se dá jednoduše spočítat jako [1]

$$v = (331,6 + 0,61 \{t\}) \text{ m} \cdot \text{s}^{-1}. \quad (1.1)$$

V mé práci budu počítat s hodnotou  $340 \text{ m} \cdot \text{s}^{-1}$ .

### 1.1.2 Akustický tlak

Je brán jako rozdíl mezi okamžitou výchylkou celkového tlaku a vztažnou referenční hodnotou  $p_0$ , což je barometrický (atmosferický) tlak. Kvůli kmitání vlny (pro zjednodušení můžeme vzít funkci sinus) se tedy tlak harmonicky mění v čase za přítomnosti zvuku. Tyto výchylky poté můžeme zachytit na mikrofón, což je ve své podstatě měnič výchylky akustického tlaku na elektrickou vlnu (1.1).

### 1.1.3 Akustická impedance

Šíření tlaku (zvuku) má za následek rozkmitání prostředí. Kinetická energie částic a potenciální energie prostředí se přelévá, což se dá analogicky popsat jako pohyb nábojů v elektrickém poli. Pro definování akustické impedance (speciální akustické impedance na jednotkové ploše) je odvozen vzorec:[1]

$$z_a = \frac{p}{v} = \rho \cdot c_0, \quad (1.2)$$

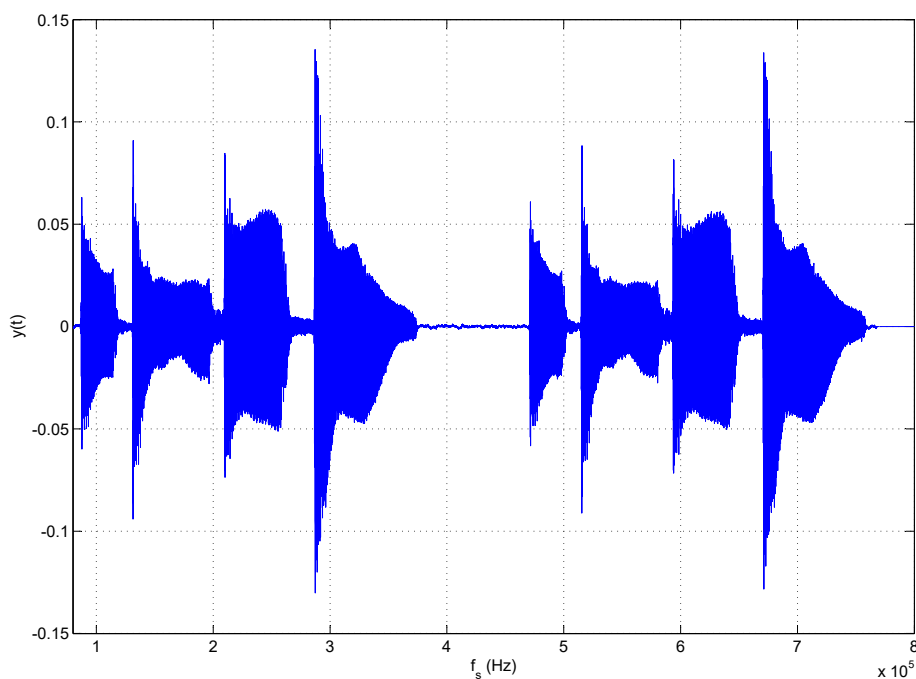
kde  $p$  je akustický tlak v místě a  $v$  je akustická rychlost. Díky akustické impedanci se na rozhraní dvou prostředí může zvuk lámat nebo odrážet. Jeho velikost závisí na mnoha veličinách (atmosferický tlak, teplota vzduchu, hustota prostředí). V mé práci budu brát velikost akustické impedance vzduchu  $z_A = 413 \text{ kg} \cdot \text{m}^2 \cdot \text{s}^{-1}$ . [1]



### 1.1.4 Typy zvukových polí

Jak již bylo naznačeno, existuje tzv. volné pole, kde neexistují odrazy. Pokud je ve směru zvuku nějaká překážka, musí zákonitě dojít k odrazu a pohlcení energie, případně k ohybu zvukové vlny. V uzavřeném prostoru poté dochází k interferenci odražených (sekundárních) a přímých (primárních) vln a pokud se tento děj opakuje, tak akustické pole odražených vln nazýváme difúzní pole. V difúzním poli nedokážeme určit směr šíření zvukové energie, jelikož je náhodný.

Za předpokladu, že sekundární vlny v difúzním poli mají malou energii, můžeme toto pole zanedbat a považovat prostor za volné pole. Tento předpoklad jsem využil právě při pokusech v bezodrazové komoře. Dalším rozřazením je vzdálené pole. Ve vzdáleném poli vzdálenost zdroje zvuku od mikrofону či pole mikrofónů je srovnatelná s velikostí tohoto zdroje. Potom se vlna považuje za rovinnou a nemusí se brát v potaz zakřivení vlnoplochy.[1]



Obr. 1.1: Ukázka nahraného zvuku z laboratoře bez vnějšího ruchu.

## 1.2 Lokalizace směru přicházejícího zvuku

Předpokládejme bodový zdroj zvuku a dva dokonale přenášející mikrofony, které jsou ve vzdáleném poli. Pro lokalizaci směru zvuku potřebujeme zjistit úhel, pod kterým dopadá na vybraný mikrofón z mikrofonního páru. Existuje celá řada metod jak určit směr přicházejícího zvuku např.:

1. pomocí směrovosti
2. pomocí rozdílné energie
3. pomocí časového zpoždění dopadu na mikrofón.

V této práci se budu hlavně věnovat určování polohy pomocí časového zpoždění.

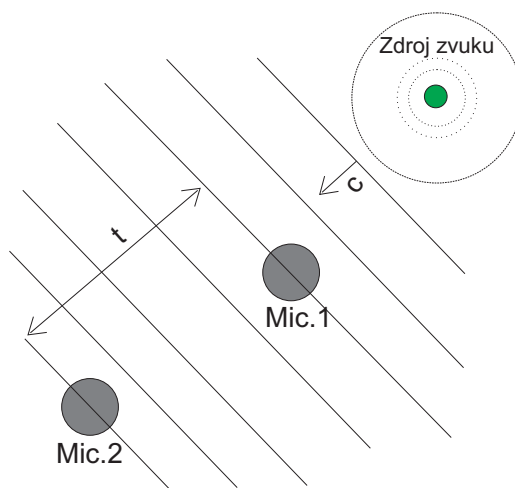
### 1.2.1 Metody pro určování polohy zvuku pomocí časového zpoždění

Metoda určování polohy pomocí časového zpoždění je ze všech metod 1.2 nejjednodušší a dostatečně přesná, a přesto i zde máme hned několik postupů. [2]

Můžeme použít:

1. všeobecná křížová korelace (GCC – Generalized Cross Correlation),
2. transformaci fáze (PHAT – Phase Transform),
3. pravděpodobnostní maximum (ML – Maximum Likelihood)

a další. Nejzákladnějším algoritmem je křížová korelace (anglicky cross-correlation), na které jsou jednotlivé postupy založené. A právě proto je tato práce vystavěna právě na ní a zmíním se o ní v samostatné kapitole.



Obr. 1.2: Časové zpoždění mezi dvěma mikrofony ve volném poli.

## Odhad časového zpoždění

Pro matematické vyjádření odhadu budu předpokládat, že zvukový zdroj je bodový. Potom můžeme zpožděný signál na dvou mikrofonech  $x_1, x_2$  popsat funkcemi [4] [2]

$$\begin{aligned}x_1(t) &= s_1(t) + n_1(t) \\x_2 &= \alpha s_1(t + d) + n_2(t),\end{aligned}\tag{1.3}$$

kde  $s_1(t)$  je zvukový signál ze zdroje,  $n_1(t), n_2(t)$  je Gaussovský, nezávislý šum prostředí přidaný do signálu.  $\alpha$  je útlum energie signálu a  $d$  je časový rozdíl. Tato rovnice nezahrnuje dozvuk prostředí, což by mohlo vést k nepřesným výsledkům. Přidání dozvuku můžeme realizovat pomocí konvoluce [4], [2]

$$x_i(t) = g_i * s(t) + n_i(t),\tag{1.4}$$

kde  $g_i$  je impulzní odezva daného prostoru,  $i$  v indexu představuje pořadové číslo mikrofону,  $s(t)$  vyjadřuje zvukový zdroj,  $n_i(t)$  značí šum,  $*$  je znaménko konvoluce.

## Odhad směru zvuku pomocí 2 mikrofónů

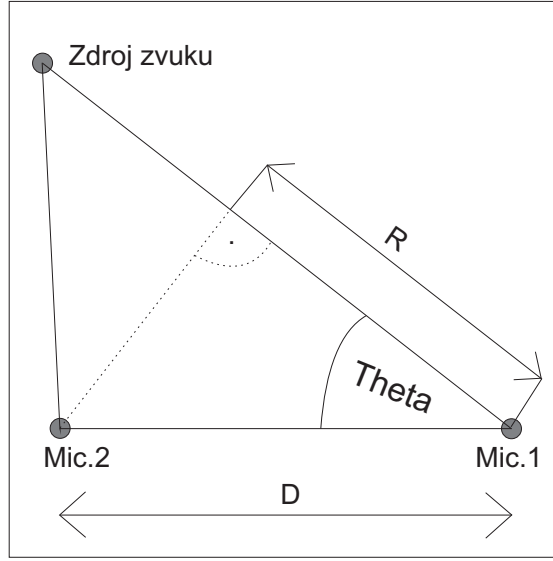
Pomocí dvou mikrofónů (ležící v prostoru na osách  $x, y$ ) dokážeme lehce určit úhel přicházejícího zvuku, avšak již nedokážeme určit, jestli zvukový zdroj leží mimo plochu  $x, y$  (jestli jde zvuk zdola či shora). Abychom dokázali určit posun na ose  $z$ , tak mikrofony musí ležet na osách  $x, y, z$ , z toho plyne, že je nutné použít minimálně tři mikrofony.

Z obrázku (1.3) je patrné, že neřešíme zatím prostředí s dozvukem. Je možné do signálu uměle přidat šum, který napodobí alespoň trochu reálné podmínky.

Úhel jednoduše podle obrázku můžeme spočítat jako [2], [3]

$$\theta = \arccos \frac{\Delta t \cdot c}{d}.\tag{1.5}$$

Proměnná  $d$  je vzdálenost mezi dvěma měřícími mikrofony,  $\Delta t$  je časové zpoždění zvuku mezi mikrofony a  $c$  je rychlost zvuku při dané teplotě vzduchu (vlhkost a tlak zanedbávám). Jsou i metody, které využívají více mikrofónů, pro mé účely však zatím budou stačit dva.



Obr. 1.3: Schéma zvuku dopadajícího na dva mikrofony a náčrt úhlu dopadajícího zvuku.

### Křížová korelace

Je hned několik metod, jak zjistit časový rozdíl, avšak nejjednodušší metodou je křížová korelace. [2] Jestliže známe dva průběhy signálů, z nichž jeden z nich je zpožděný, potom nalezením maxima jejich křížové korelace získáme časový posun signálu mezi mikrofony. Zobecněním rovnic (1.3) pro více signálů vyplývá tvar [2]

$$x_i = a_i s_i(t + \tau_i) + n_i(t), \quad (1.6)$$

kde  $i$  v indexu značí, ze kterého mikrofonu je brán signál  $s(t)$ ,  $\tau_i$  je hledaný časový rozdíl mezi zvolenými mikrofony,  $a_i$  označuje útlum signálu. V rovnici je přidán šum  $n(t)$ , který musí být na signálu nezávislý, jinak by mohl být považován za signál, což by mohlo vést ke znevážení výsledků. Podrobněji se touto závislostí zabývá G. C. Carter [3]. A právě kvůli šumu je v mnoha metodách hledání  $\tau_i$  (transformace fáze, pravděpodobnostní maximum) signál filtrován váhovací funkcí. Abychom mohli snáze najít  $\tau_i$ , je nutné provést Fourierovu transformaci signálu, kterou více popisují v sekci matematických funkcí(1.3). Získáme rovnici ve tvaru [2]

$$X_i(f) = \alpha_i S_i(f) \cdot e^{-i2\pi f \tau_i} + N_i(f), \quad (1.7)$$

kde  $f$  je frekvence. Potom výsledná obecná korelační funkce (křížová výkonová hustota spektra) vyjde [2], [4]

$$R_{x_1 x_2} = E[x_1(t) \cdot x_2(t - \tau)]. \quad (1.8)$$

$E[.]$  je předpokládaná funkce, která se liší podle toho, jak chceme signál filtrovat. Pokud za  $E = 1$ , potom dostaneme křížovou korelaci v časové doméně. Příkladem jiné funkce může být třeba tato rovnice [4]

$$R_{x_1x_2}(t) = \frac{1}{\tau - T} \int_{\tau}^T x_1(t) \cdot x_2(t - \tau) dt. \quad (1.9)$$

### Všeobecná křížová korelace

Všeobecná křížová korelace je vylepšená verze křížové korelace. Největší výhodou tohoto algoritmu je relativní přesnost a malé nároky na výpočetní funkci. Jak už bylo zmíněno, nejčastěji se používá signál, který již prošel přes číslicový filtr, a to kvůli tomu, aby výsledná korelace měla jednoznačné maximum a byla zajištěna vyšší přesnost výsledku. Kromě číslicových filtrů tuto skutečnost zajišťuje například lepší odstup signálu od šumu (SRN), správně zvolený vzorkovací kmitočet, analogové filtry vřazené do signálové cesty mezi mikrofon a samotný proces nahrávání. [2], [4], [3]

$$S_{x_1x_2} = E[X_1(f) \cdot X_2^*(f)] \quad (1.10)$$

je výkonová spektrální hustota (určuje výkon signálu odpovídající frekvenčnímu pásmu) mezi signály  $x_1, x_2$ , které již prošly Fourierovou transformací. Předpokládaná funkce  $E[.]$  má stejný význam jako v (1.8) a  $(.)^*$  značí komplexně sdružené číslo. Rovnice všeobecné křížové korelace v tomto případě potom bude [2]

$$\Psi_{x_1x_2}(f) = \int_{-\infty}^{\infty} \phi(f) S_{x_1x_2}(f) e^{i2\pi f\tau} dt. \quad (1.11)$$

V této rovnici  $\phi(f)$  značí kmitočtovou váhovací funkci, což může být například impulzní odezva prostoru. Integrál v praxi nahrazujeme zpětnou Fourierovou transformací. Časové zpoždění  $\tau_\phi$  získáme tak, že najdeme maximum této rovnice, matematicky zapsáno [2], [4], [3]

$$\tau_\phi = \arg \max \Psi_{x_1x_2}(f). \quad (1.12)$$

Výběrem váhovacího filtru v rovnici (1.11) značně ovlivníme výsledek a je jasné, že čím složitější funkci zvolíme, tím větší nároky klademe na výpočetní techniku. Pro obecnou křížovou korelaci bychom zvolili  $\phi(f) = 1$ . Další metody určování časového rozdílu, například transformace fáze nebo pravděpodobnostní maximum, jsou odvozeny právě z křížové korelace a liší se pouze ve váhovacím filtru. Pouhá křížová korelace je velice účinná v prostředí, kde není velký dozvuk (menší než 0,15 s) a signál není rušen šumem či hlukem. [2]

## Pravděpodobnostní maximum

Pravděpodobnostní maximum je metoda založená na křížové korelaci, získáme ji z rovnice (1.11) dosazením tohoto filtru [4]

$$\psi_{\text{ML}}(f) = \frac{1}{|G_{x_1x_2}(f)|} \cdot \frac{|\gamma_{x_1x_2}(f)|^2}{1 - |\gamma_{x_1x_2}(f)|^2}, \quad (1.13)$$

kde

$$R_{\text{ML}}(\tau) = \int_{-\infty}^{\infty} \psi_{\text{ML}}(f) \cdot X_1(f) \cdot X_2^*(f) \cdot e^{i2\pi f\tau} dt = \int_{-\infty}^{\infty} \psi_{\text{ML}}(f) \cdot G_{x_1x_2}(f) \cdot e^{i2\pi f\tau} dt, \quad (1.14)$$

$\psi_{\text{ML}}$  je váhovací funkcí,  $X_i(f)$  značí Fourierovu transformaci  $x_i(f)$ ,  $X_1^*(f)$  je komplexně sdružené číslo,  $G_{x_1x_2}(f)$  představuje výkonové křížové spektrum přijímaného signálu. Abychom mohli spočítat pravděpodobnostní maximum (jeho váhovací funkci (1.13)), zjišťujeme míru koherence vstupních dat, což znamená, do jaké míry může vycházet signál  $x_1$  ze signálu  $x_2$ . [4], [2]

$$|\gamma_{x_1x_2}(f)|^2 = \frac{|G_{x_1x_2}(f)|^2}{G_{x_1x_1}(f)G_{x_2x_2}(f)}. \quad (1.15)$$

Rovnice využívá metodou nejmenších čtverců, což oproti PHAT (1.2.1) by mělo být přesnější, protože křížové spektrum není pouhou konstantou, ale snaží se jej dle nějaké křivky kopírovat. Časové zpoždění můžeme zjistit pomocí maxima [2]

$$\hat{\tau}_{\text{ML}} = \arg \max_{\tau \in D} R_{\text{ML}}(\tau). \quad (1.16)$$

Tato metoda je díky tomu, že korelaci provádí ve spektrálních složkách (kde je lepší odstup signálu od šumu), velmi přesná. Její nedostatek je v tom, že na určitých frekvencích interferuje a může tak tvořit lokální maxima. [7]

## Transformace fáze

Základem této metody je, že vezme křížové spektrum signálů  $x_1, x_2$  a převede jej na konstantu. Váhovací funkci transformace fáze můžeme vyjádřit jako [2]

$$\psi_p(f) = \frac{1}{|G_{x_1x_2}(f)|}, \quad (1.17)$$

dosazením do (1.11) potom získáme [2]

$$R_p(\tau) = \int_{-\infty}^{\infty} \frac{1}{|G_{x_1x_2}(f)|} G_{x_1x_2}(f) e^{j\omega\tau} dt. \quad (1.18)$$

Výsledné časové zpoždění vyjádříme ve tvaru [2]

$$\hat{\tau}_p = \arg \max_{\tau \in D} R_p(\tau). \quad (1.19)$$

Tato funkce zachovává fázi signálů, proto se dá použít i na kontrolování polarity. Hlavní výhodou je, že vylučuje z křížového spektra energii signálu, která v ostatních filtrech může ovlivnit výsledek.

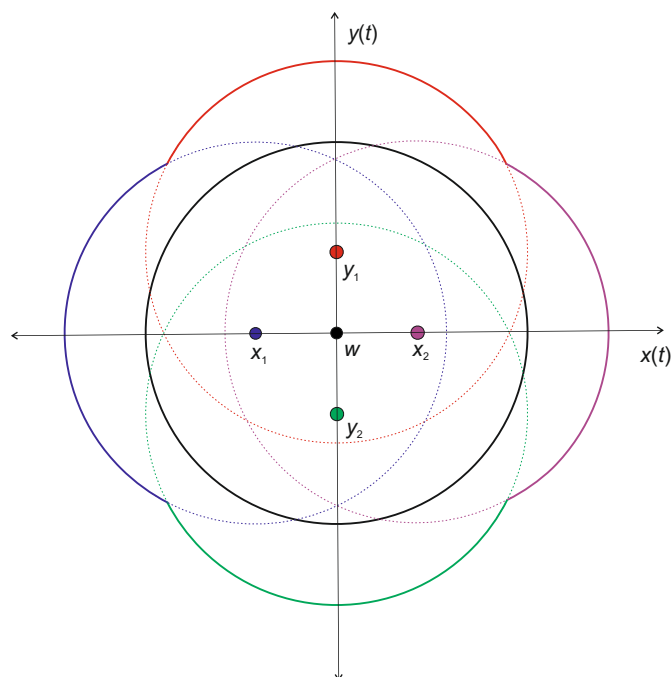
### 1.2.2 Určování polohy zvuku pomocí analýzy energického spektra

#### B-formát signálu

Tato metoda využívá B-formát signálu, který lze získat pomocí dvou mikrofónů 1. řádu (tzv. osmičkové) a jednoho mikrofónu 0. řádu (všesměrového). Dva mikrofóny 1. řádu můžeme nahradit čtyřmi mikrofóny všesměrovými, potom signál musíme ale odečíst  $p_{celkove} = p_{n_1} - p_{n_2}$ , kde  $p_1$  a  $p_2$  je tlak na jednotlivé membrány, tudíž jejich odečtení nám dá ve výsledku tlak  $\sum p$  na osmičkovém mikrofónu. [1] V tomto formátu můžeme určit, jak horizontální, tak i vertikální úhel. Samotný formát obsahuje čtyři signály:  $x(t)$ ,  $y(t)$ ,  $z(t)$ ,  $w(t)$ , ze kterých lze následně spočítat směr přicházejícího zvuku. Signály  $x(t)$  a  $y(t)$  nám dávají informace o horizontální rovině, jsou zaznamenávány právě na dva osmičkové mikrofóny (čtyři všesměrové).  $x(t)$  určuje předozadní rovinu a  $y(t)$  určuje levo-pravou rovinu. Signál  $z(t)$  reprezentuje vertikální osu a  $w(t)$  je referenční signál nahraný na všesměrový mikrofón, který však není nutný. Z goniometrie můžeme určit tento druh mikrofónu ze čtyřech mikrofónů jako  $w = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ , kde  $x_1, x_2$  jsou dva mikrofóny ležící na ose  $x$  a  $y_1, y_2$  jsou dva mikrofóny ležící na ose  $y$ . Při vzdálenosti mikrofónů od sebe zhruba do 7 mm se bude tento součet mikrofónů chovat přibližně jako jeden kulový mikrofón (za předpokladu, že všechny použité mikrofóny jsou kulové a důležité informace v signálu jsou přibližně do 12 kHz, od stoupající frekvence se začnou mikrofóny odečítat kvůli malé vlnové délce zvuku). Při zvolení této metody (ze čtyř mikrofónů spočítáme jeden) je nutné zamyslet se nad tím, jestli odmocněním neztratíme podstatné informace o fázi apod. Jednotlivé rovnice pro definování B-formát signálu potom vypadají následovně [9]:

$$\begin{aligned}x(t) &= \cos(\alpha) \cos(\beta) s(t), \\y(t) &= \sin(\alpha) \cos(\beta) s(t), \\z(t) &= \sin(\beta) s(t), \\w(t) &= \frac{1}{\sqrt{2}} s(t),\end{aligned}\tag{1.20}$$

kde  $\alpha$  je úhel v horizontální rovině a  $\beta$  je elevace zdroje zvuku oproti poli ve vertikální rovině,  $s(t)$  značí nahraný signál. Z obrázku je patrné, že ideálního B-formát signálu nikdy nedosáhneme, a to kvůli rozměrům mikrofónů.



Obr. 1.4: Ukázka nastavení mikrofونů pro B-formát. Jednotlivé dvojice leží na osách a pátý mikrofون ve středu souřadnic. Všechny mikrofony jsou kulového charakteru.

### STFT a energická metoda

Hlavním předpokladem pro tuto metodu je, že směr umístění zvuku je v opačném směru než intenzita vektoru zvuku. Vzhledem k tomu, že energie zvuku je rozložená ve spektru a čase, je dobré spektrum zvuku rozdělit do frekvenčních úseků (anglicky „bin“). K tomu slouží STFT (short time Fourier Transform), která pomocí okna rozdělí signál a následně provede Fourierovu transformaci. Typy oken můžou být různé, např. pravoúhlé, Kaiserovo nebo Hanningovo okno a další. Hlavním parametrem okna je jeho velikost (size). Pro občasné, nestacionární signály by mělo mít okno malou velikost, aby algoritmus pracoval správně. Kvůli nedokonalosti některých oken je nutné zadat ještě STFT overlapping (překryv), jinak hrozí ztráta dat u specifických signálů. [10] [11] Po provedení STFT můžeme matematicky zapsat vektor intenzity pro každý úsek takto:[9]



$$\begin{aligned}
I_x(t, f) &= \frac{1}{\sqrt{2}z_0} \operatorname{Re} \{W^*(t, f) \cdot X(t, f)\} \\
I_y(t, f) &= \frac{1}{\sqrt{2}z_0} \operatorname{Re} \{W^*(t, f) \cdot Y(t, f)\} \\
I_z(t, f) &= \frac{1}{\sqrt{2}z_0} \operatorname{Re} \{W^*(t, f) \cdot Z(t, f)\}.
\end{aligned} \tag{1.21}$$

V těchto rovnicích  $z_0$  je akustickou impedancí vzduchu,  $t$  je čas,  $f$  je frekvence daného úseku,  $(.)^*$  značí komplexně sdružené číslo,  $X(t, f)$ ,  $Y(t, f)$ ,  $Z(t, f)$  a  $W(t, f)$  jsou signály  $x(t)$ ,  $y(t)$ ,  $z(t)$  a  $w(t)$ , na které byla již aplikována Fourierova transformace. Jednotlivé vektory ukazují opačný směr přicházejícího zvuku v jednotlivých osách, takže je nutné je obrátit. Výsledný úhel v horizontální rovině potom můžeme získat pomocí:[9]

$$\alpha(t, f) = \tan^{-1} \left[ \frac{-I_y(t, f)}{-I_x(t, f)} \right], \tag{1.22}$$

pro hodnoty  $I_y \geq 0$ . Pro hodnoty  $I_y < 0$  potom platí tato rovnice:[9]

$$\alpha(t, f) = \tan^{-1} \left[ \frac{-I_y(t, f)}{-I_x(t, f)} \right] - 180^\circ. \tag{1.23}$$

Pokud bychom chtěli získat elevaci (úhel ve vertikální rovině), potom použijeme rovnici:[9]

$$\beta(t, f) = \tan^{-1} \left[ \frac{-I_z(t, f)}{\sqrt{I_x(t, f)^2 + I_y(t, f)^2}} \right]. \tag{1.24}$$

Po spočítání úhlu každého frekvenčního úseku v čase, je nutné určit úhel celkový. Tento algoritmus předpokládá, že v každém frekvenčním úseku bude dominantní pouze jeden zdroj zvuku, což jistě klade velký důraz na dobrý odstup signálu od šumu (SNR). Dále může být problém se zvuky s podobným spektrem, jelikož mohou být reprezentovány jako jeden zdroj, zvláště, pokud se rozhodneme počítat pouze v určitém frekvenčním spektru pro urychlení a zjednodušení. Pokud budeme počítat s celým spektrem a jedním zdrojem, potom je nejčastěji zastoupený úhel ten hledaný. Při více zdrojích signálu toto však nemusí platit, avšak hledaný signál by měl být v některých frekvenčních úsecích nejčtenější. Tato metoda se jeví jako dobrá pro detekci pohybu signálu s předem známým frekvenčním spektrem. Matematický zápis pro hledání výsledného úhlu je následující:[9]

$$N(\alpha) = \sum_{k=0}^K p(S(t, f_k)|\alpha), \tag{1.25}$$

kde  $K$  je počet frekvenčních úseků pro  $\alpha \in (-\pi, \pi)$ ,  $S$  je zvukový signál,  $t$  značí index časového úseku,  $f$  je frekvenční úsek a  $p(S(t, f_k)|\alpha)$  je pravděpodobnost úhlu  $\alpha$

přicházejícího signálu. Po spočtení všech pravděpodobností úhlu vybereme celkový nejpravděpodobnější úhel  $\alpha_{\text{vysledna}}$  přicházejícího zvuku takto: [9]

$$\alpha_{\text{vysledna}} = \arg \max N(\alpha). \quad (1.26)$$

## 1.3 Matematické funkce

### 1.3.1 Diskrétní Fourierova transformace (DFT)

Další důležitou funkcí je DFT (Discrete Fourier Transform), která převádí signál z časové oblasti do frekvenčního spektra složek. Než je možné provést DFT, je nutné na vstup funkce poslat diskrétní signál, nemůžeme použít spojitou funkci. V praxi se používá vzorkování (sampling), které nám uloží hodnoty signálu  $s(t)$  v určitých diskrétních časových intervalech daných vzorkovací frekvencí (sample rate).

Navzorkovaný signál musí splňovat Nyquistův (Shannonův) teorém (1.27), který říká, že vzorkovací frekvence musí být minimálně dvakrát větší než nejvyšší frekvence harmonických složek obsažených ve vzorkovaném signálu [6]

$$2f_{\text{vz}} = f_{\text{max}}, \quad (1.27)$$

kde  $f_{\text{vz}}$  je vzorkovací frekvence a  $f_{\text{max}}$  maximální frekvence harmonických složek. Pokud tento teorém není dodržen, potom dochází k aliasingu (překryvu kmitočtů vzorkovaného signálu), což je nežádoucí, často se proto používá antialiasingový filtr, který má za úkol odstranit kmitočty vyšší než  $\frac{f_{\text{vz}}}{2}$ . Příkladem může být nahrávání signálu, který chceme korelovat. Kdybychom podcenili vzorkovací frekvenci, mohlo by dojít k zašumění signálu a šum by byl navíc závislý, což by značně ovlivnilo výsledek, jak popisují v 1.3.4. Aliasingu si můžeme všimnout ve více oborech, v hudebním zpracování, filmu nebo třeba v grafice.

Diskrétní Fourierova transformace pracuje s konečným signálem o  $N$  vzorcích v diskrétním čase, narozdíl od spojitě Fourierovy transformace. Proto je nutné přepsat integrál ve spojitěm čase (1.28) [6]

$$s(t) = \int_{-\infty}^{\infty} \hat{s}(f) \cdot e^{i2\pi ft} dt, \quad (1.28)$$

na sumu diskrétních vzorků

$$\hat{s}_n = \sum_{k=0}^{N-1} s_n \cdot e^{-i2\pi ik \frac{n}{N}} = \sum_{k=0}^{N-1} s_n \cdot W_N^{kn}. \quad (1.29)$$

Amplitudové spektrum DFT je symetrické podle osy  $y$ , stejně jako je u spojitě Fourierovy transformace.

### 1.3.2 Rychlá Fourierova transformace (FFT)

FFT je algoritmus pro vypočítání DFT. Její hlavní výhodou je rychlost, kdy oproti klasickému postupu DFT, dokáže pro velký počet vzorků být rychlejší o cca.  $10^4$  vzorků, samozřejmě závisí na metodě FFT (tento údaj je pro FFT, která používá „decimaci v čase“). [8]

#### Využití

Rychlá Fourierova transformace má mnoho využití. Nejčastěji se s ní setkáme právě při úpravě či zjišťování dat ze zvukového signálu. Může se jednat o konvoluci, korelaci, zobrazení frekvenčního spektra zvuku, ale i u různých digitálních filtrů. Ve zpracování obrazu se uplatňuje speciální dvourozměrná transformace, je také možné pomocí FFT dopočítat fraktálovou texturu (uplatnění zejména u počítačových her). V neposlední řadě se může použít pro matematickou aritmetiku (například dopočítání čísla  $\pi$  na miliontiny). [8]

### 1.3.3 Krátkodobá Fourierova transformace (STFT)

Krátkodobá Fourierova transformace (Short-Time Fourier Transform) je algoritmus, který je velice podobný FFT, akorát zpracovává signál po částech, které jsou vymezeny tzv. oknem.

#### STFT v diskrétním čase

Vstupním signálem je již navzorkovaný signál v čase, což provádí A/D převodník. Každý vzorek v čase je rozdělen na několik pásem pomocí okna a je převeden do frekvenční oblasti pomocí FFT. Matematicky lze STFT napsat jako:[12]

$$\text{STFT} \{x[n]\}(m, \omega) \equiv X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n] \cdot w[n-m] \cdot e^{-i\omega n}, \quad (1.30)$$

kde  $x[n]$  je vstupní signál a  $w[n]$  je okno, kterým se signál násobí. Výraz  $e^{-i\omega n}$  značí Fourierovu transformaci, v praxi se potom často zavádí rychlá Fourierova transformace. Pokud je potřeba zmenšit okno  $w$  k dosažení lepších výsledků, může se posunout o parametr  $m$ , kterým by měla být nespojitá hodnota. [12]

#### Vybraná okna pro STFT

Okna pro STFT mohou být různá, např. obdélníkové, Bartlettovo, Welchovo, Hanningovo, Hammingovo, Blackmann–Harrisovo, či Nuttallovo. Podrobný přehled a zhodnocení je k nalezení v referenčním zdroji [11]. Já se zmíním pouze o pár oknech, které jsou vhodné pro algoritmus energetické metody. Jen nutné si uvědomit,

že čím složitější okno, tím větší bude nárok na výkon počítače a také je nutné znát, k jakému účelu okna používáme, jestli špatně zvolené okno může ovlivnit výsledky měření.

### **Hanningovo okno**

Hanningovo okno je jedno z oken, která jsou založena na váhovacím součtu konstantního výrazu a kosinovou funkcí výrazu. Zapsat jej můžeme jako [11]

$$z = \frac{2\pi j}{N}, \quad j = 0 \dots N - 1, \quad (1.31)$$

$$w_j = \frac{1 - \cos(z)}{2} = \cos^2\left(\frac{z - \pi}{2}\right). \quad (1.32)$$

Hanningovo okno má nízké kmitočtové ztráty, a proto se používá v mnoha případech, kde nevadí zkreslení spektra sinusové vlny (například u měření šumu).

### **Blackman–Harrisovo okno**

Základ algoritmu je ve sčítání kosinové funkce. Existuje více druhů tohoto okna, proto popíši jeden vzorový příklad. Jde o takzvané „92 dB Blackman–Harrisovo okno“, které bylo nakonfigurováno speciálně tak, aby boční laloky v přenosové funkci, které sousedí s hlavním, byly malé.[11]

$$z = \frac{2\pi j}{N}, \quad j = 0 \dots N - 1, \quad (1.33)$$

$$w_j = 0,35875 - 0,48829 \cos(z) + 0,14128 \cos(2z) - 0,01168 \cos(3z), \quad (1.34)$$

takto vypadá matematické zaznamenání. Blackman–Harrisovo okno má velice nízké kmitočtové ztráty a také velice dobrou chybovost v šířce pásma a amplitudě. Díky malým postranním lalokům se hodí i na detekci malých sinusových signálů, které jsou kmitočtově blízké velkým signálům. Například pokud Fourierovou transformací dostaneme velkou amplitudu základní frekvence a k tomu malé amplitudy dalších složených frekvencí. Další výhodou může být, že funkce má velký dynamický rozsah.

### 1.3.4 Křížová korelace

Hlavní matematickou funkcí pro určení směru je křížová korelace, která udává podobnost signálu vůči druhému (speciální případ je autokorelace, podobnost signálu sobě samému). Vstupem jsou dvě funkce, výstupem je jedna. Matematická definice je [5]

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} x(t)y(t + \tau)dt, \quad (1.35)$$

Maximum této funkce je v bodě, kdy jsou si vstupní signály  $x(t)$  a  $y(t + \tau)$  nejvíce podobné, v nejlepším případě autokorelace stejné. Minima bude rovnice 1.35 dosahovat tehdy, když vstupní funkce budou mít shodné tvary, ale jedna z nich bude kladná a druhá záporná (budou v protifázi). Potom se výsledky budou blížit nule.

Z popisu výše je patrné, že přidaný šum do signálu popisovaný v rovnici (1.3) musí být náhodný a nezávislý na vstupním signálu.[4] Také právě proto je výhodné převádět vstup do frekvenční roviny Fourierovou transformací, neboť potom můžeme přímo porovnávat jednotlivé složky signálu a ne celý, těžko bychom porovnávaly dva signály v časové doméně.

### Využití

Nejznámějším využitím korelační funkce je v sonaru nebo radaru. Zařízení vyšle signál do prostoru, přijímá zašuměný odražený a z následného zpoždění dokážeme určit tvar překážky, od které se vlna odrazila. Ale v praxi má tato funkce daleko širší použití. Jedná se například o analýzu a rozpoznávání biologických objektů v obraze nebo akustické zaměřování objektu.

Mikrofonní pole, je sada mikrofónů, která je postavena nejčastěji do geometrických tvarů. Mikrofonní pole dělíme podle polohy v prostoru na:

1. lineární,
2. rovinné,
3. prostorové.

V lineárním poli jsou mikrofony dány do jedné přímky (osa  $x$ ), nevýhodou je možnost určení polohy pouze v jednom směru. Do rovinného pole se mohou mikrofony rozmístit v ploše, a tudíž musí být ve stejné výšce (osy  $x, y$ ). Posledním tvarem je pole prostorové, ve kterém nejsou mikrofony v žádném tvaru, leží volně v prostoru (osy  $x, y, z$ ).

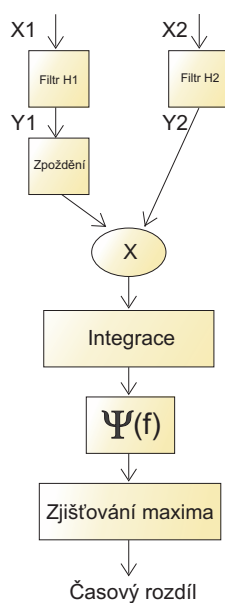
Mikrofonní pole podle tvaru seskupení:

1. čtvercové,
2. kruhové,
3. křížové,
4. kvadrátové,
5. sférické.

Každé pole má své výhody i nevýhody. Největší rozdíly jsou v následném správném určení úhlu, odkud přichází zvuk.

## 1.4 Matlab

Název pochází z anglického MATrix LABoratory, což již naznačuje, že pracuje především s maticemi. Je zde také možnost symbolického zápisu (například pro matematické rovnice), této možnosti jsem však ve své práci nevyužil. Hlavním grafickým prostředím je příkazové okno, které je tvořeno příkazovým řádkem sloužící k výpočtům, spouštění skriptů atd. V Matlabu podporuje možnost psát si své vlastní funkce do skriptů. Novější verze mohou mít knihovny s předprogramovanými funkcemi, které ve starších verzích nejsou a nefungují.



Obr. 1.5: Schéma GCC korelace

## 2 VYHODNOCENÍ PRAKTICKÉ ČÁSTI

V této sekci budu popisovat praktické uskutečnění teoretických poznatků výše. Jako programovací prostředí byl použit simulační program Matlab, snímání nahraných zvuků bylo realizované v bezodrazové komoře nebo byly uměle vytvořeny. K veškerým měřením bylo použito mikrofونů AUDIX TM1, což jsou testovací a měřicí kulové mikrofony, které jsou dostupné v laboratoři FEKT.

### 2.1 Popis měření a zpracování výsledků

Nahrávací vzorkovací kmitočet jsem zvolil 48 kHz, přičemž je jasné, že čím větší vzorkovací kmitočet bude, tím bude i přesnější měření. Na druhou stranu nutno podotknout, že při vyšším vzorkování bude i větší velikost souboru a větší náročnost na zpracování dat. Můžeme si totiž lehce odvodit vzorec, jaký může být největší rozdíl ve vzorcích  $\Delta s_n$  mezi dvěma mikrofony na vzdálenosti  $d$ :

$$\Delta s_n = \frac{f_s d}{c}, \quad (2.1)$$

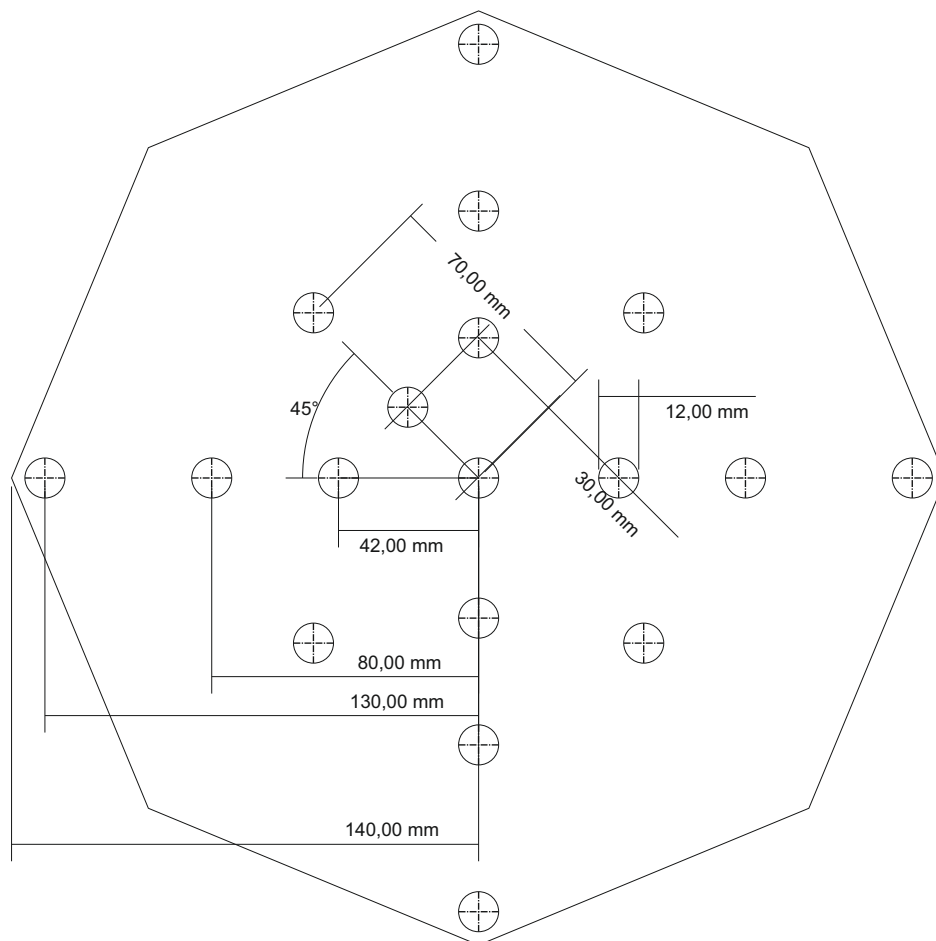
kde  $f_s$  je vzorkovací frekvence,  $d$  vzdálenost mezi mikrofony a  $c$  je rychlost zvuku. Tento vzorec vychází z tabulek níže (2.1 2.2). Z toho plyne, že na větší vzdálenost dokážeme zaznamenat větší zpoždění, samozřejmě musíme stále předpokládat, že mikrofony jsou ve volném poli. Je také nutné správně určit rychlost zvuku v daném prostředí, protože na vzdálenosti 5 cm může mikrofون zaznamenat (při sámplovací frekvenci 48 kHz) 6 vzorků při rychlosti zvuku  $345 \text{ m} \cdot \text{s}^{-1}$  a 7 vzorků při rychlosti  $340 \text{ m} \cdot \text{s}^{-1}$ . A na počtu vzorků je přímo úměrné dělení úhlů, tzn. rozmezí  $0-90^\circ$  bude při 7 vzorcích rozděleno na 7 částí, které ale nebudou stejně velké (kvůli funkci  $\arccos$ ).

#### 2.1.1 Parametry měření bez vyrobeného mikrofonního pole

Pro hledání směru zvuku v malém mikrofonním poli jsem zvolil křížové pole o osmi mikrofonech, jak je vidět z obrázku 2.2. Malé mikrofonní pole je definováno do 20 cm v průměru, tudíž jsem jednotlivé dvojice mikrofونů na stojanu rozestavil přibližně od sebe 5 cm. Uprostřed tedy zbyla mezera 10 cm. Zvuk jsem nahrával v bezodrazové místnosti s dozvukem  $RT60 < 0,3 \text{ s}$ . Jako zkušební signál jsem použil nahrávky struny kytary a cinknutí zvonku kvůli jejich frekvenčním předpokladům (nízké a vysoké kmitočty v nahrávce). Pohyb zvuku jsem simuloval pomocí reproduktoru, který jsem vždy po přehrání jedné nahrávky posunul. Celkem jsem nastavil 6 poloh zvuku, v prvním kvadrantu 4 pokusy (tj. na  $0^\circ$ ,  $30^\circ$ ,  $60^\circ$  a  $90^\circ$ ), potom další dvě polohy na  $180^\circ$  a  $270^\circ$ .

## 2.1.2 Parametry měření s vyrobeným mikrofonním polem

### Návrh a realizace držáku na mikrofonní pole



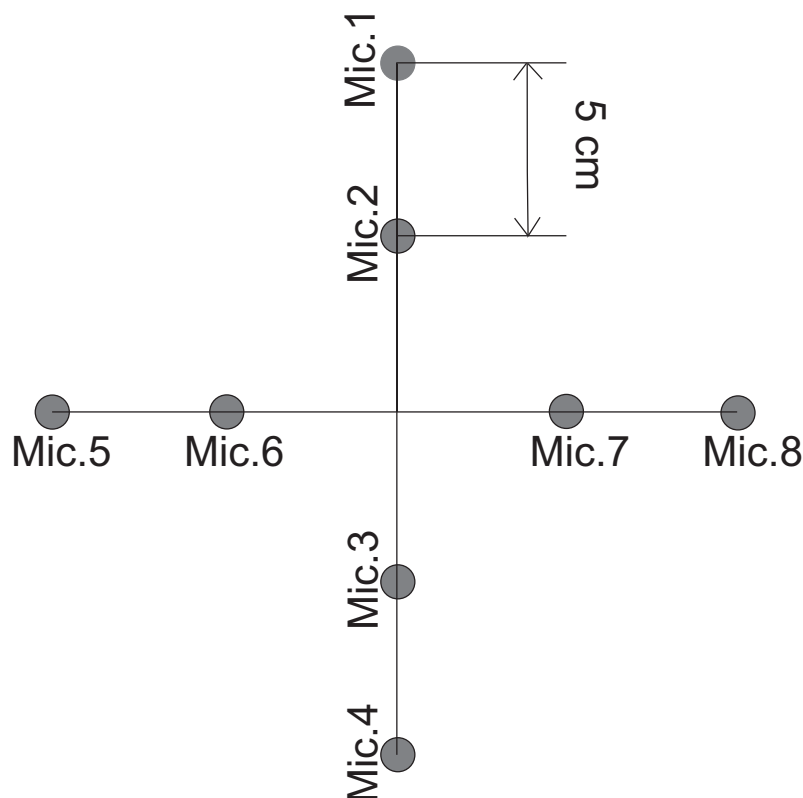
Obr. 2.1: Navržené mikrofonní pole

Abych zlepšil výsledky měření, a zároveň zjednodušil instalaci mikrofonů, tak jsem navrhl držák mikrofonního pole. Lze na něj instalovat hned několik typů mikrofonních polí, křížové, kruhové a pole pro měření energetickou metodou. Otvor uprostřed slouží k zachycení na stojan, ostatní pro mikrofony. Jednotlivé parametry (vzdálenosti děr) jsou změřeny a navrženy speciálně pro měřicí mikrofony, průměr děr je navržen pro klasické držáky. Křížové pole je popsáno již v grafice 2.2.

Kruhové pole má průměr 20 cm. Při energetické metodě jsou mikrofony uchycovány tak, aby byly co nejbližší u sebe. Původní návrh držáku byl kruhový a z plechu. Nakonec pro jednodušší výrobu jsem volil osmiúhelník a materiálem je plast o tloušťce 3,6 mm. Držák na pole byl vyroben a ozkoušen. Pro zlepšení bych nechal více prostoru pro kabely mikrofonu, protože u tohoto návrhu je nutné použít zalomené XLR

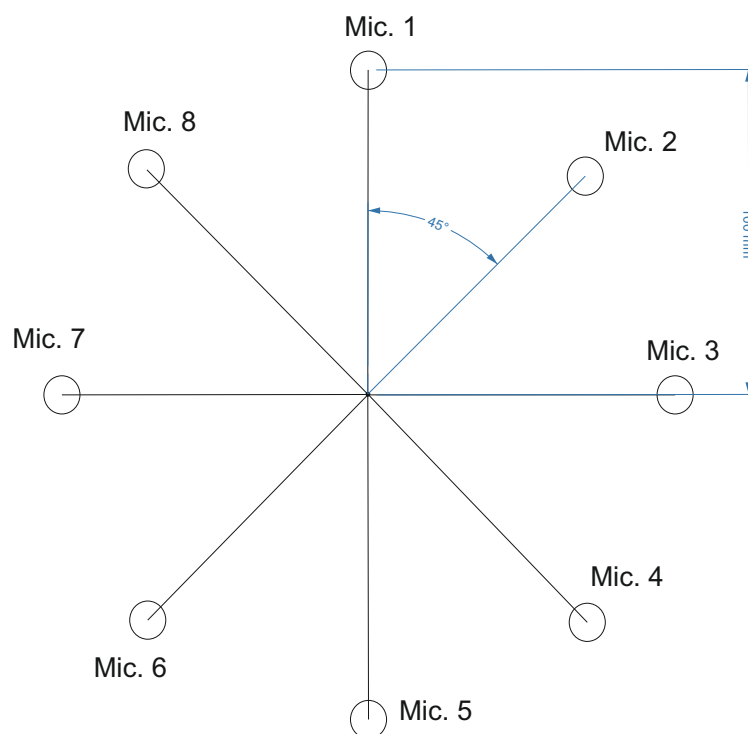


kabely. Co je nutné dále řešit je prokluzování držáků samotných mikrofonů, nicméně stačí pro řešení tohoto problému použít podložky. Jinak mikrofony jsou ve správných vzdálenostech i výšce, takže hlavní účel držák plní bez problémů, snad u kruhového pole se špatně odhaduje poloha mikrofonů ležících v úhlu  $45^\circ$  od středu, ale zde pomáhá zvolený tvar osmiúhelníku držáku.



Obr. 2.2: Schéma rozestavení mikrofonů v křížovém poli.

Jako pokusný signál jsem volil bouchnutí dveří, tříštění skla a výstřel z pistole, a to opět kvůli jejich frekvenčnímu spektru. Pohyb zvuku byl simulován pomocí natáčení pole pomocí rotátoru, který je dostupný v laboratoři FEKT. Rotátor je spojen s dálkovým ovládáním, jde na něm nastavit přesný úhel, o kolik se mikrofonní pole posune. Z předem vypracovaných výpočtů jsem se rozhodl rotovat pole v intervalu  $\langle 0^\circ, 90^\circ \rangle$ , protože v intervalu úhlů  $(90^\circ, 360^\circ)$  by měl být signál stejný, ale s jinými znaménky.



Obr. 2.3: Schéma rozestavení mikrofónů v kruhovém poli.

### 2.1.3 Zpracování v Matlabu

V této části se zabývám postupem, který jsem naprogramoval. Nejdříve je nutné nahrané vzorky (2.1) nahrát do proměnných. Jelikož se jedná již o diskretní signál, tak Matlab vytvoří matici o  $1 \times N$  polí, kde  $N$  je počet samplů. Pro další zpracování je tento formát nevhodný, tudíž je lepší z něj udělat matici o  $N \times 1$  polích, což se provede inverzí matice.

```
[y,fs] = audioread ();
Y=[y, y1,y2,y3,y4,y5,y6,y7];
Y=Y';
y= Y(1,:);
```

Audioread je příkaz k načtení zvukového souboru do proměnné  $y$  se vzorkovací frekvencí  $fs$ . Nyní můžeme přistoupit k samotnému zjišťování časového zpoždění. Funkce FFT, IFFT (zpětná rychlá Fourierova transformace) jsou v Matlabu již naprogramované, a proto jsem jich využil. Pro správný výstup křížové korelace musí mít vstupní signál stejnou délku, což je třeba ošetřit, například stejnou časovou

délkou nahranych vzorků nebo uměle zarovnat matici. Následně už můžeme provést FFT obou signálů ze vstupu a vyjádříme z nich výkonovou spektrální hustotu, kterou potom vydělíme její velikostí (absolutní hodnotou). Když takto vydělenou výkonovou spektrální hustotu převedeme zpět do časové osy pomocí IFFT, dostaneme korelační funkci, jejíž maximum značí časovou prodlevu mezi jednotlivými signály na vstupu. Pro zjednodušení ovládání jsem vytvořil funkci, která provádí výše popsany algoritmus, ve tvaru:

```
function posun = faze(x,x1,d,fs),
```

kde  $x$  je signál z jednoho mikrofону,  $x1$  signál druhého mikrofону,  $d$  je vzdálenost mikrofónů v metrech,  $fs$  je vzorkovací frekvence. Pro různé metody jsou vstupní parametry stejné  $x, x1, d, fs$ , takže jsem funkce jednotlivých metod jenom jinak pojmenoval. Výstupními parametry jsou:

```
plot(fx);
disp(['index maxima: ' num2str(index_maxima)]);
disp(['posun [s]: ' num2str(posun)]);
disp(['úhel [stupeň]: ' num2str(theta)]);
```

Jako první výstup je graf korelace, index maxima dává informaci, na kolikátém vzorku se největší shoda signálů nachází. Posun, jehož jednotkou jsou sekundy, udává o kolik jsou signály časově posunuty a poslední parametr úhel ukazuje úhel od určeného mikrofónu ke zdroji ve stupních. Celý kód funkce metody GCC je v příloze A.1 a jednotlivé další metody v přílohách A.2 a A.3.

Při zjišťování posunu signálu musíme dát pozor, aby

$$\Delta t < \frac{d}{c}, \quad (2.2)$$

kde  $\Delta t$  je časové zpoždění,  $d$  vzdálenost mezi mikrofóny,  $c$  značí rychlost zvuku. Je to nutné kvůli tomu, že funkce arccos platí v reálných číslech pouze na intervalu  $\langle -1, 1 \rangle$ . Pokud tuto podmínku nedodržíme, Matlab vypíše výsledek v komplexním čísle, což není žádoucí, protože nedokážeme vypočítat zpoždění, dále je tato hodnota ignorována.

Pokud nemáme nahrané reálné vzorky zvuku, potom lze vytvořit v Matlabu matici s šumem, která nám nahradí matici se zvukem a následné zjišťování zvukového zpoždění je obdobné jak u nahrávek. Stačí do jedné ze dvou matic vložit na začátek nuly, což vlastně znamená, že mikrofón nepřijímá signál. Do druhé matice je nutno vložit stejný počet nul, akorát na konec. Každé políčko matice vlastně znamená 1 vzorek mikrofónu, z čehož vyplývá, že v mém případě pro správnost výsledků nemůžu posunout šum o více než 7 samplů. Podrobněji zmíněno v sekci 2.1.

Abych mohl jednoduše zpracovat výsledky z celého pole najednou, vytvořil jsem si jednoduchou funkci, která sama spočítá pohyb předmětu.

```
function pohyb = ALL_methods(X, Y, pole, metoda, fs, c)
```

Matice se zvukem musí být již zpracované tak, aby vstupní matice měla v řádcích jednotlivé mikrofony v poli. Vstupy jsou X, Y. Jsou zde zahrnuty metody GCC (generalized cross-correlation), PHAT (phase transform) a ML (maximum likelihood). X je první poloha předmětu (nahraný zvuk mikrofonním polem) a Y je druhá poloha, v matici může být až osm mikrofónů. Volby: pole = 0 je křížové pole, pole = 1 je kruhové pole; metoda = 0 je GCC, metoda = 1 je PHAT, metoda = 2 je ML. Proměnná  $fs$  je vzorkovací frekvence v Hz a  $c$  je rychlost zvuku v m/s. Nyní ukáži část algoritmu výpočtu pohybu zvukového zdroje pro celé pole:

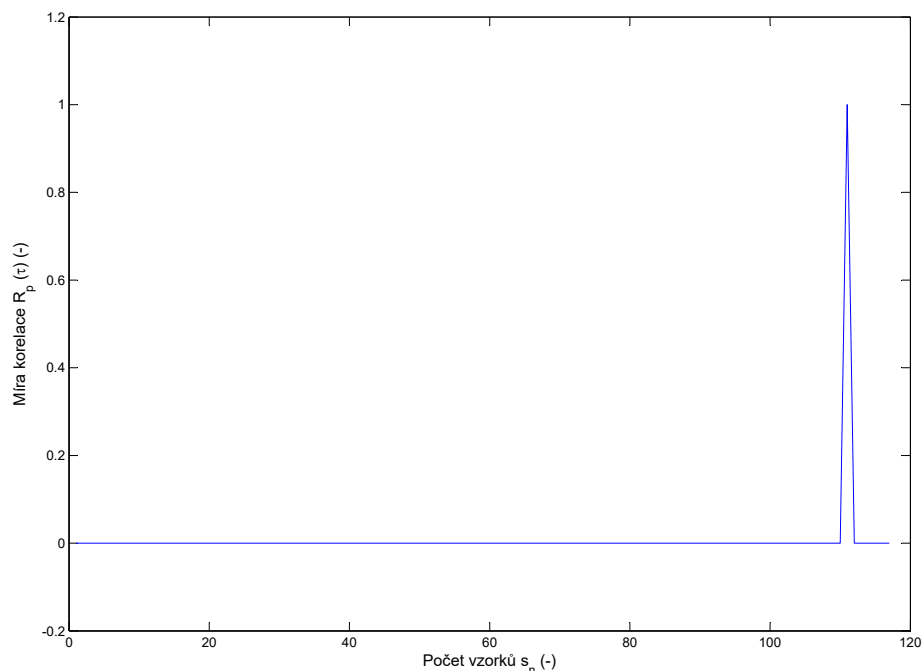
```
for i = 1:8
    for j = (i+1):8
        x = Y(i,:);
        x1 = Y(j,:);
        vzdalenost = d(i,j);
        theta(i,j) = acosd(posun*(c/vzdalenost));
        if imag(theta(i,j)) ~= 0
            theta(i,j) = abs(theta(i,j));
        end
    end
end
```

Zjednodušeně jde o cyklus, který propočítá jednotlivé výsledné úhly mezi mikrofony v poli. Nadefinoval jsem matici se vzdálenostmi mezi jednotlivými mikrofony, tudíž algoritmus bere vždy mikrofony na prvku (i,j) a stejně vezme vzdálenost na prvku (i,j). Jde vlastně o souřadnice. Výslednou matici úhlů potom ještě kompenzují odečtením úhlů natočení mikrofónů na stejném principu souřadnic. Podmínka pro imaginární čísla působí to, že pokud je nějaký výsledek v imaginární oblasti, tak se z něj udělá absolutní hodnota, což vlastně je jeho velikost.

## 2.2 Výsledky

### 2.2.1 Výsledky zkušebního signálu

V této kapitole jsou shrnuty výsledky měření popsaného výše. Zpracování signálu probíhalo jak je popsáno v 2.1.3. V tabulce (2.1) jsou výsledky časového posunu ( $\Delta t$ ) a odpovídající úhel ( $\theta$ ) dopadajícího naprogramovaného šumu, jehož korelaci můžeme vidět v grafu (2.4).



Obr. 2.4: Graf korelovaného naprogramovaného zvuku

Maximum funkce je velice dobře zřetelné, protože se jedná o nasimulovaný nerušený signál v bezodrazovém prostředí. Počítal jsem výsledky pro vzdálenost mikrofónů  $d = 5, 10$  cm a  $f_s = 48,96$  kHz. Z tabulky lze vyčíst, že pro  $d = 5$  cm čím větší časový rozdíl v rozmezí  $\langle 0; 1,4583 \cdot 10^{-4} \rangle$  s, tím víc nelineárně roste úhel, což je dáno funkcí  $\arccos$ . Průměrná změna úhlu v rozmezí zpoždění od 0 do 5 vzorků cca.  $9^\circ$  a v oblasti od 5 do 7 vzorků  $18,752^\circ$ . V praxi to znamená, že je nutné snímat více mikrofony než dvěma, aby byla chyba u větších časových rozmezí co nejmenší.

Tabulka (2.2) popisuje časová zpoždění téhož šumu, akorát snímány dvojnásobnou vzorkovací frekvencí.

Jde vidět, že při dvojnásobné vzorkovací frekvenci bude přesnost určení úhlu dvakrát větší, jelikož se měřený interval rozdělí na dvakrát více vzorků. Na druhou stranu, čím vyšší kmitočet zvolíme, tím větší nároky klademe na parametry nahrávacího přístroje.

Co se týče vzdálenosti mikrofónů, tak z tabulek je možno vyčíst, že má na výsledek podobný vliv jako vzorkovací kmitočet, tzn. pokud vezmeme 5 cm rozestup, dostaneme maximálně 7 vzorků. Při 10 cm rozestupu je to 14 vzorků atd. Opět se nám dvakrát zlepší výsledky měření. Je však nutné brát v potaz i praktickou stránku jako je sestavení pole, jestli se bude jednat o pevnou instalaci nebo přenosnou.

Tab. 2.1: Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 48 kHz.

Index maxima	$\theta$ [°]	$\Delta t$ [s]
0	90	0
1	98,1443	$2,0833 \cdot 10^{-5}$
2	106,4592	$4,1667 \cdot 10^{-5}$
3	115,1507	$6,25 \cdot 10^{-5}$
4	124,5181	$8,333 \cdot 10^{-5}$
5	135,0995	$1,0417 \cdot 10^{-4}$
6	148,2117	$1,25 \cdot 10^{-4}$
7	172,598	$1,4583 \cdot 10^{-4}$
8	180-29,268i	$1,666 \cdot 10^{-4}$

Tyto dvě části měření prakticky dotvrzují vzorec 2.2. Pro měření z této simulace plyne, že použiji spíše pole kruhové, s referenčním mikrofonom uprostřed. Dále z tabulek vyplývá, že je nutné před měřením „zkalibrovat“ mikrofonní pole tak, abych dokázal správně určit úhel. Provedení bude takové, že do místa, které určím jako 0°, postavím zdroj zvuku a potom všechny mikrofony „vynuluji“ tím, že k nim přičtu nebo odečtu právě naměřený úhel.

### 2.2.2 Výsledky s reálným signálem bez vyrobeného pole

Moje práce s reálným zvukovým vzorkem probíhala dříve, než jsem naprogramoval testovací kód, což se podepsalo i na výsledcích. Při určování úhlu jsem si nejdříve zvolil špatnou část signálu, který byl zašuměný, tudíž algoritmus nepracoval správně (zde se potvrdila teorie o přidaném šumu v sekci 1.3.4, graf (2.5) ukazuje křížovou korelaci špatně zvoleného úseku signálu).

V grafu (2.6) můžeme vidět výsledek v podobě výsledného grafu obecné křížové korelace signálu. Výsledek je oproti jiným metodám velice nepřesný, předpoklad, že tato metoda nebude vhodná pro určování úhlu v prostředí s dozvukem, je docela jasné už i z těchto výsledků. Můžeme si všimnout, že jsou zde dvě velká maxima, která by mohla značně ztížit hledání úhlu.

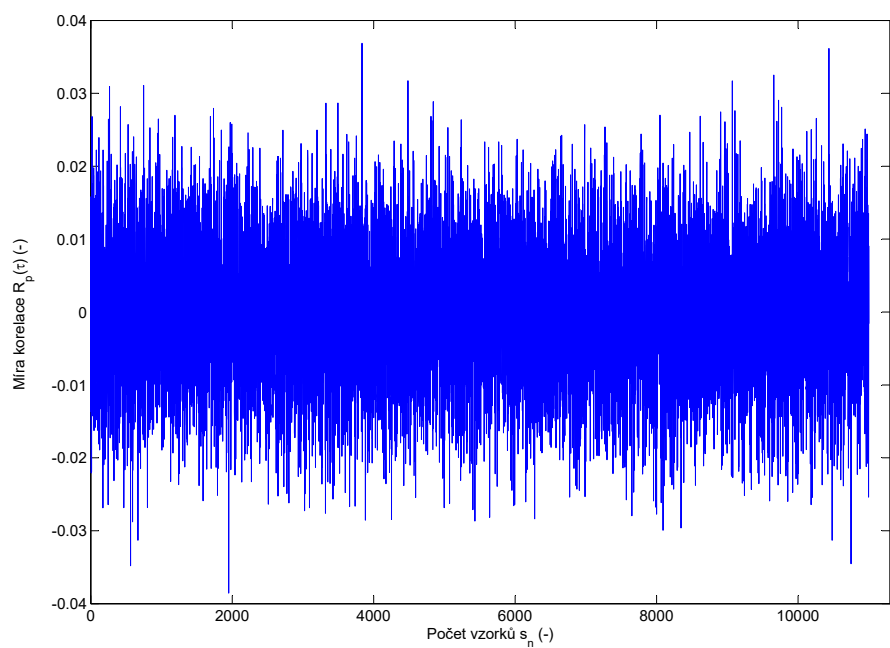
Další metoda, kterou jsem na hledání úhlu použil, byla tzv. transformace fáze a jeví se jako dobrá varianta pro měření v terénu. Na obrázku korelace metodou PHAT (2.7) vidíme nějaká lokální maxima, avšak to, které hledáme, je několikrát větší. Je to dáno tím, že používáme již filtr, který odstraní nepotřebné složky signálu.

Třetí metoda pravděpodobnostní maximum (Maximum Likelihood–ML) je z popisovaných metod nejsložitější na naprogramování, ale nedosahuje tak přesných vý-

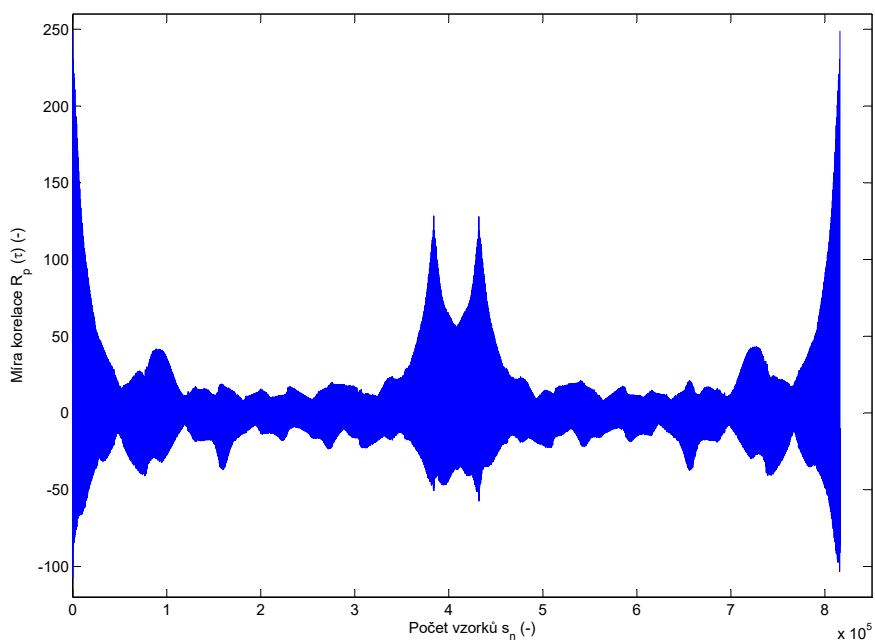
Tab. 2.2: Posun úhlu v závislosti na zpoždění při vzorkovacím kmitočtu 96 kHz.

Index maxima	$\theta$ [°]	$\Delta t$ [s]
0	90	0
1	94,0619	$1,04 \cdot 10^{-5}$
2	98,1443	$2,0833 \cdot 10^{-5}$
3	102,2689	$3,14 \cdot 10^{-5}$
4	106,4592	$4,1667 \cdot 10^{-5}$
5	110,7424	$5,2083 \cdot 10^{-5}$
6	115,1507	$6,25 \cdot 10^{-5}$
7	119,7247	$7,2917 \cdot 10^{-5}$
8	124,5181	$8,333 \cdot 10^{-5}$
9	129,6057	$9,375 \cdot 10^{-5}$
10	135,0995	$1,0417 \cdot 10^{-4}$
11	141,1843	$1,1458 \cdot 10^{-4}$
12	148,2117	$1,25 \cdot 10^{-4}$
10	157,0482	$1,3542 \cdot 10^{-4}$
14	172,598	$1,4583 \cdot 10^{-4}$
15	180-20,1531i	$1,5625 \cdot 10^{-4}$

sledků jako PHAT, co se týče naprogramovaného zvuku i nahraného. U nahraného zvuku je graf korelace ML totožný s PHAT. Avšak oproti obecné korelaci je v jejím grafu mnohem méně lokálních špiček, takže je přesnější. Výslednou korelaci můžeme vidět na obrázku (2.8).

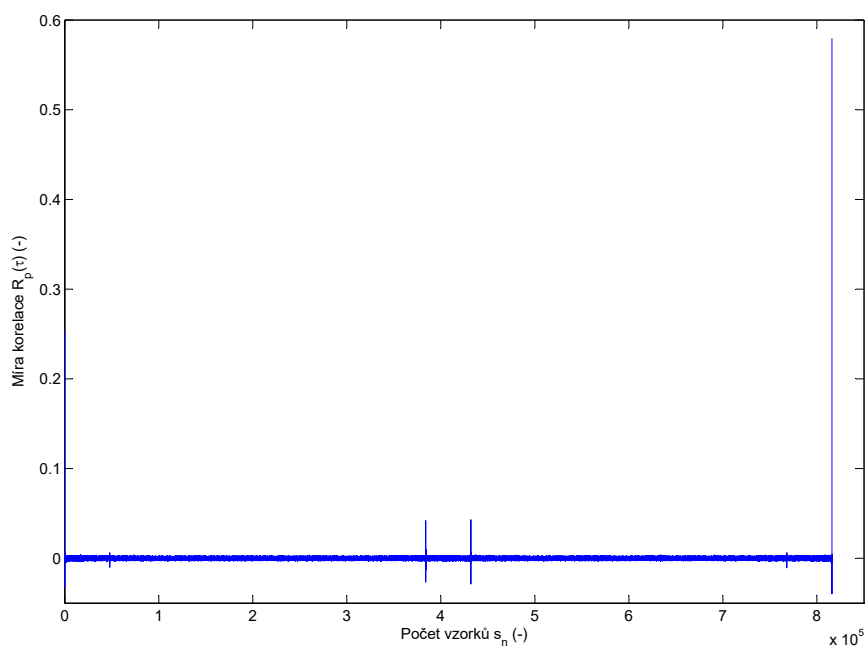


Obr. 2.5: Korelace zašuměného signálu

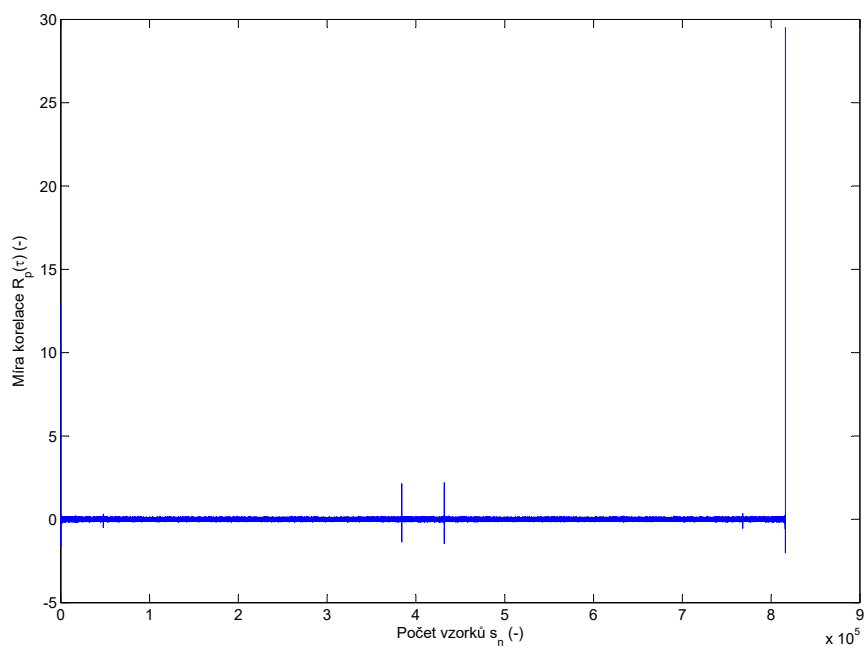


Obr. 2.6: Graf obecné křížové korelace





Obr. 2.7: Graf korelace metodou PHAT



Obr. 2.8: Graf korelace metodou ML

## 2.2.3 Výsledky s reálným signálem a s vyrobeným polem

### Výsledky z měření v laboratoři

Popis průběhu měření je v kapitole Popis měření a zpracování výsledků 2.1. Pro hledání směru zvuku v malém mikrofonním poli jsem zvolil křížové a kruhové pole o osmi mikrofonech, dále potom B-formát mikrofonů pro energetickou metodu. Jednotlivé nákresy polí je vidět z obrázků 2.2 pro křížové pole, pro kruhové pole 2.3, pro B-formát 1.4. Zvuk jsem nahrával v bezodrazové místnosti s dozvukem  $RT60 < 0,3$  s, která se nachází na půdě FEKT.

Jako signály jsem použil nahrávku bouchnutí dveří, tříštění skla a výstřel z pistole kvůli jejich frekvenčním předpokladům (obsahují nízké a vysoké kmitočty v nahrávce). Pohyb zvuku jsem simuloval pomocí rotace mikrofonního pole, které bylo zasazeno do rotátoru, a které jsem vždy po přehrání jedné nahrávky posunul. Celkem jsem nastavil 8 poloh natočení pole. Pro signál zabouchnutí dveří jsem volil úhly  $0^\circ$ ,  $8^\circ$ ,  $16^\circ$ ,  $24^\circ$ ,  $34^\circ$ ,  $44^\circ$ ,  $56^\circ$ ,  $70^\circ$  a  $90^\circ$ . Pro signál tříštění skla jsem volil jiné úhly  $0^\circ$ ,  $10^\circ$ ,  $20^\circ$ ,  $30^\circ$ ,  $40^\circ$ ,  $50^\circ$ ,  $60^\circ$ ,  $70^\circ$  a  $90^\circ$ , pro zvuk výstřelu z pistole pak  $0^\circ$ ,  $6^\circ$ ,  $15^\circ$ ,  $22^\circ$ ,  $37^\circ$ ,  $47^\circ$ ,  $58^\circ$ ,  $70^\circ$  a  $85^\circ$ . Různé úhly jsem volil kvůli detailnějšímu zpracování, na druhou stranu tato volba lehce znepráhledňuje grafy a výsledky.

Příklad zjišťování úhlů směru přicházejícího zvuku můžeme vidět v tabulce 2.5. Jde vidět, že obyčejným průměrováním výsledných úhlů z mikrofonu můžeme dostat velice nepřesný výsledek, a to kvůli mikrofonům, které nevyhodnotili signál správně. Zlepšení výsledků by mohlo přinést váhování výsledků, popřípadě počítání průměru z vybraných hodnot (např. počítání dalšího průměru z hodnot kolem průměru do padesáti procent). V tabulkách 2.3 nalezneme vyhodnocení jednotlivých polí a metod. Jedná se o průměry z absolutních hodnot úhlů všech mikrofonů v poli, protože funkce, ze které počítáme úhel, při prohození vstupních signálů vypočítá stejnou hodnotu, ale s jiným znaménkem. Nejhuře z algoritmů dopadla metoda GCC 1.2.1, metody PHAT a ML v laboratoři jsou stejné, což se mohlo stát kvůli úzkému spektru zvuku (grafy jednotlivých metod mají stejný charakter jako 2.7 a 2.8). Potvrdilo se, že čím větší úhel se snažíme zjistit, tím větší chybu dostaneme. Proto je dobré měřit v malých intervalech, které nám zajistí dostatečnou přesnost. Přehled jednotlivých odchylek v závislosti na úhlu pohybu najdeme v tabulkách 2.9, 2.10, 2.11.

### Výsledky z měření v reálném prostředí

Měření v reálném prostředí proběhlo pouze s šesti mikrofony, a to kvůli nedostupnosti zvukové karty s více vstupy. Mikrofony jsem rozestavěl do křížového pole, jelikož dávalo jedny z nejlepších výsledků, s tím, že mikrofony číslo 6 a 7 2.2 jsem vynechal. Bylo k tomu nutné přizpůsobit kód v Matlabu, jelikož jinak by vznikly

rozdíly mezi velikostmi matic. Testovací signály byly: zvuk ze sirény, generátoru zvuku rozbití skla a startovací pistole. Měření probíhalo v prostředí, kde ve vzdálenosti cca. 8 metrů byla cihlová zeď, tudíž v záznamu byl silný zvuk z odrazu. Dalšími ruchovými aspekty byly: vítr, zpěv ptáků, štěkání psů, lidé atd. Jednalo se tedy o docela rušné prostředí. Teplota vzduchu byla okolo 30 °C. Správný úhel jsem odhadoval pomocí sítě dlaždic, což by mělo být relativně přesné, ale nejedná se o laboratorní přesnost, a tedy průměrné odchylky mohou být  $\pm 3$  stupně. Výsledky z nahrávání jsou zaznamenány v tabulce 2.4. V praxi se ukázalo, že metoda GCC může mít velkou chybovost, avšak relativně dobře fungovala na vzorcích, což je možná i trochu překvapující (metodou průměrování mikrofónů se GCC značně zlepšila, při jiných metodách by však výsledky dopadly podstatně hůře). U výstřelu žádná z metod nezaznamenala pohyb na 40 stupňů, což může být dáno limitací signálu v rekordéru, ale i jinými faktory.

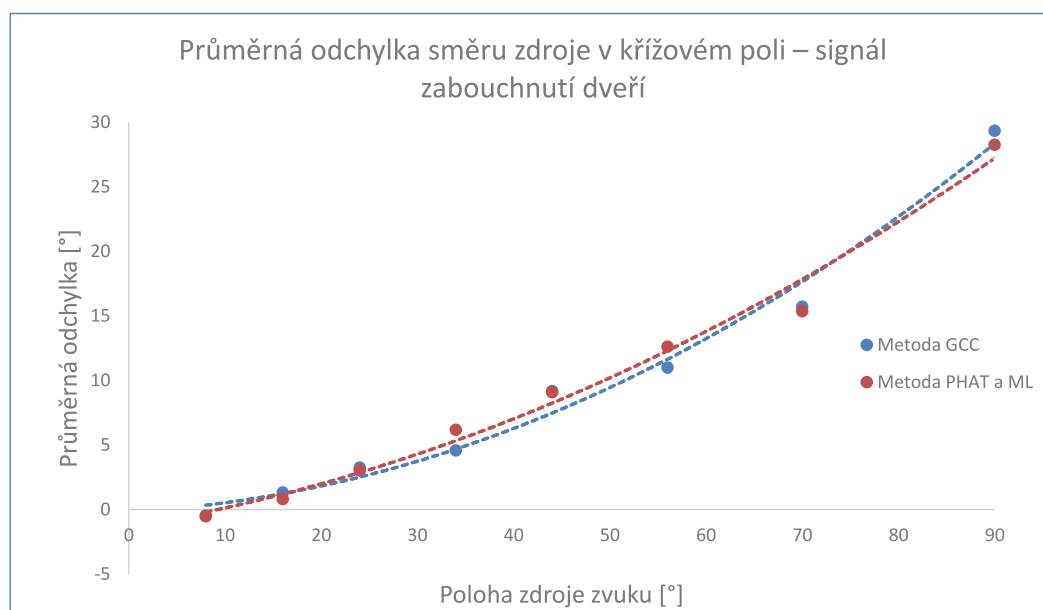
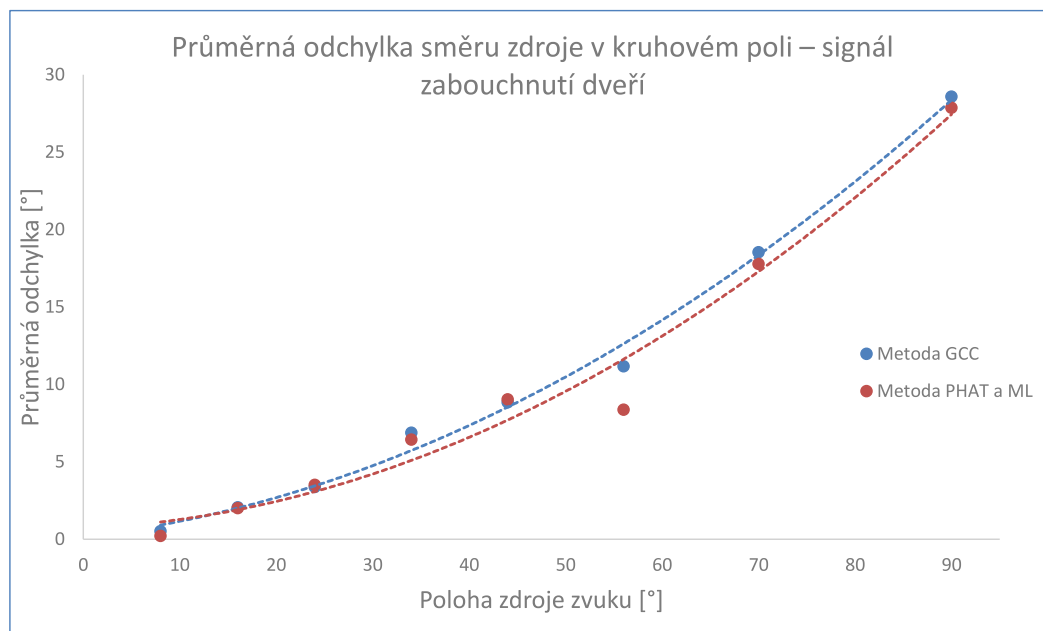
Po dohodě s vedoucím práce byla provedena čtyři měření (tři výsledky) u výstřelu z nedostatku nábojů. Při zaznamenávání zvuku sirény celý pokus narušil nepřiměřený hluk, tudíž jsem byl nucen přerušit měření a byla použita pouze čtyři měření. Měření s generátorem zvuku rozbití skla proběhlo bez problémů. Celkově však nahrávání i algoritmy splnily účel a fungují.

Tab. 2.3: Tabulka výsledků měření v laboratoři, přehled jednotlivých metod

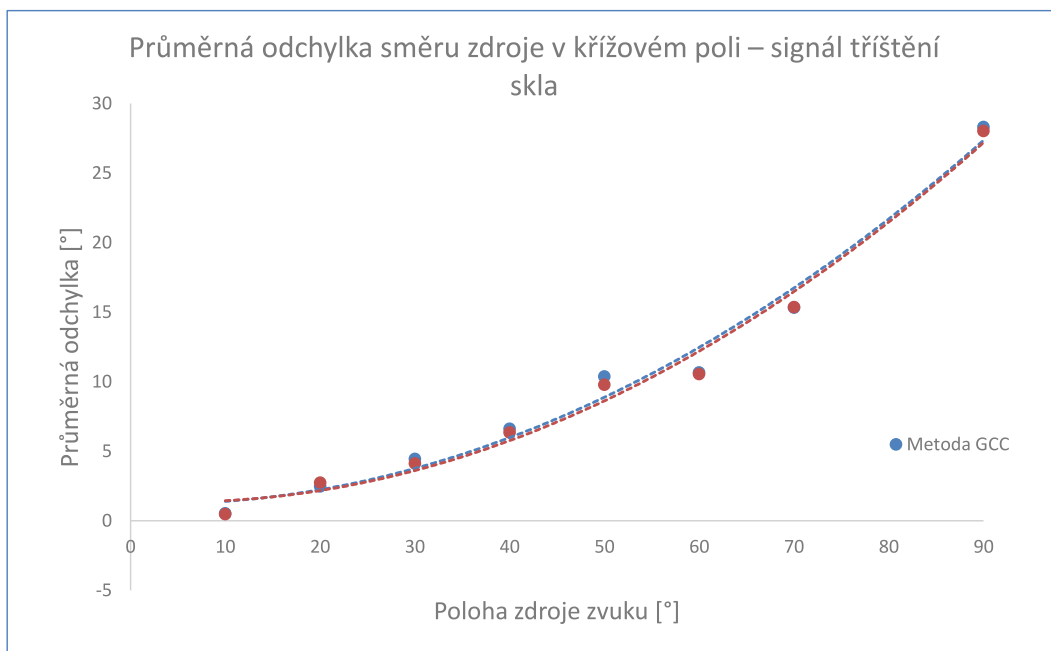
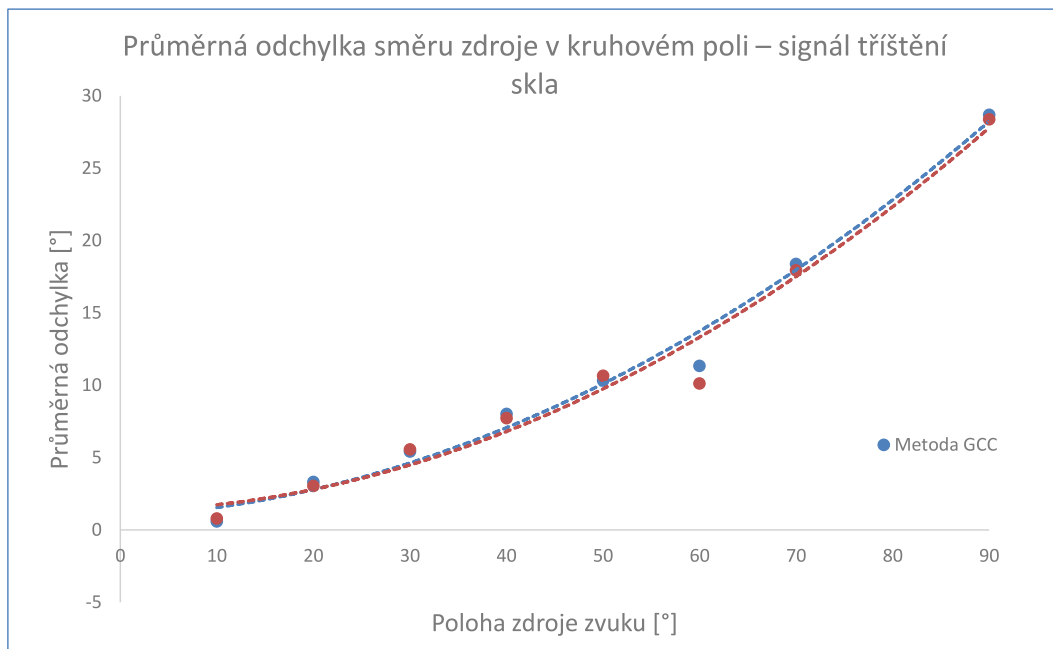
Výsledky mikrofonních polí v laboratoři						
Bouchnutí dveří						
Pohyb ve °	Kruhové pole			Křížové pole		
	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
8	7,47	7,78	7,78	8,46	8,54	8,54
16	13,94	13,98	13,98	14,69	15,20	15,20
24	20,63	20,48	20,48	20,75	20,99	20,99
34	27,12	27,56	27,56	29,43	27,83	27,83
44	35,15	34,96	34,96	34,84	34,91	34,91
56	44,83	47,64	47,64	45,01	43,38	43,38
70	51,46	52,22	52,22	54,29	54,64	54,64
90	61,41	62,13	62,13	60,66	61,74	61,74

Výsledky mikrofonních polí v laboratoři						
Rozbití skla						
Pohyb ve °	Kruhové pole			Křížové pole		
	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
10	9,42	9,20	9,20	9,46	9,52	9,52
20	16,68	16,97	16,97	17,53	17,25	17,25
30	24,57	24,44	24,44	25,54	25,86	25,86
40	31,98	32,28	32,28	33,37	33,64	33,64
50	39,68	39,34	39,34	39,61	40,21	40,21
60	48,66	49,89	49,89	49,33	49,45	49,45
70	51,62	52,06	52,06	54,66	54,63	54,63
90	61,30	61,63	61,63	61,67	61,97	61,97

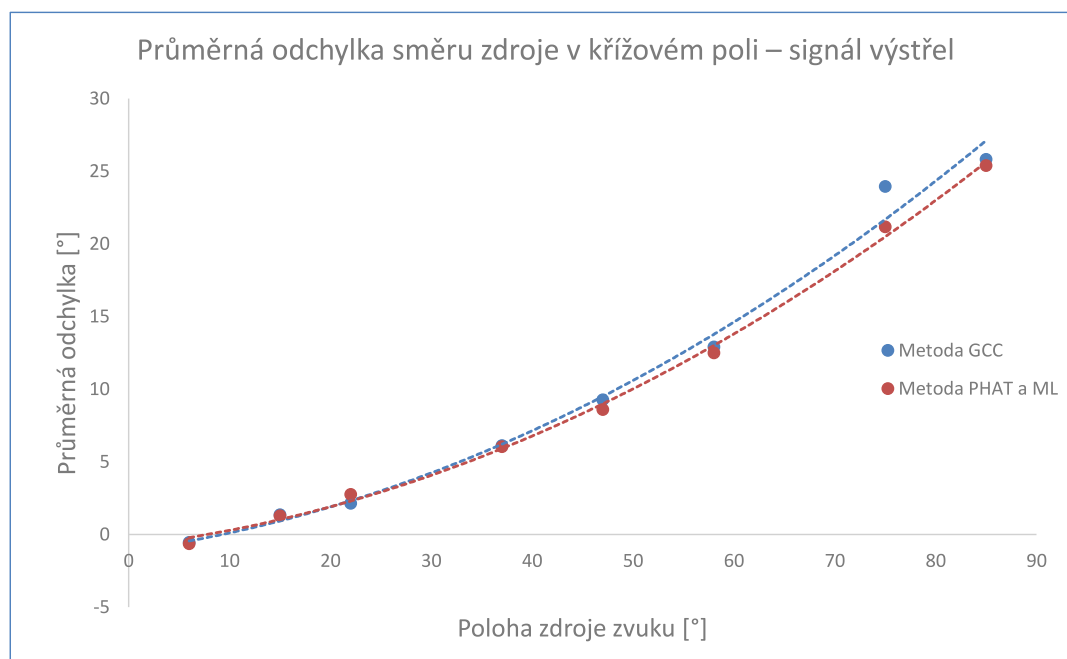
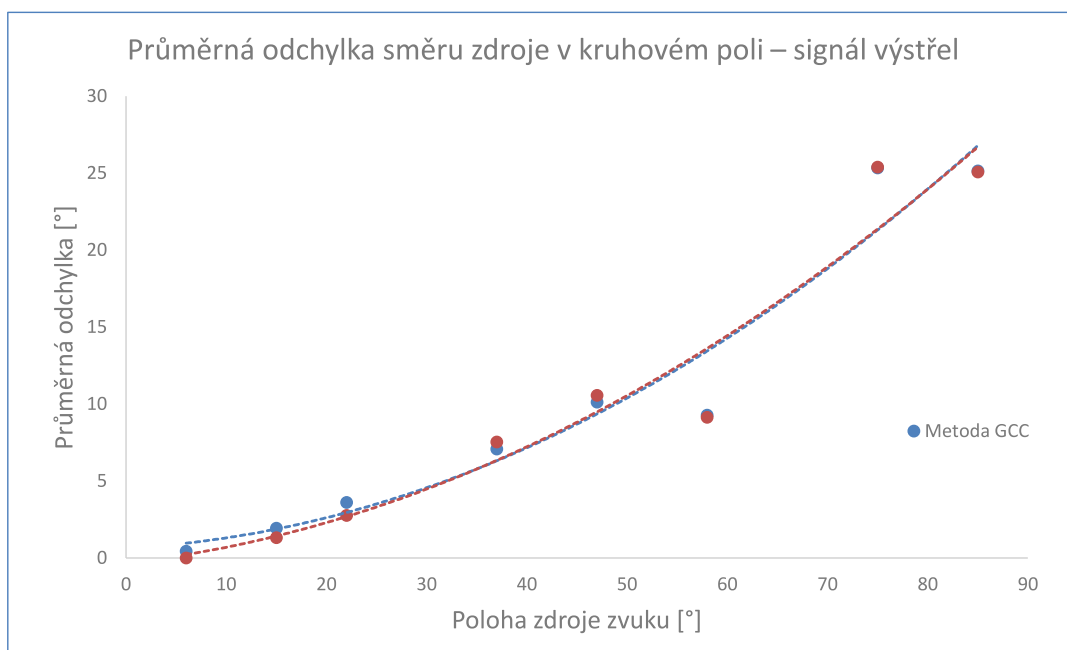
Výsledky mikrofonních polí v laboratoři						
Výstřel						
Pohyb ve °	Kruhové pole			Křížové pole		
	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
6	5,55	5,99	5,99	6,54	6,63	6,63
15	13,05	13,67	13,67	13,63	13,71	13,71
22	18,38	19,23	19,23	19,85	19,24	19,24
37	29,92	29,47	29,47	30,88	30,95	30,95
47	36,87	36,43	36,43	37,72	38,39	38,39
58	48,72	48,86	48,86	45,09	45,50	45,50
75	49,65	49,60	49,60	51,04	53,83	53,83
85	59,85	59,93	59,93	59,19	59,61	59,61



Obr. 2.9: Grafy průměrné odchylky, měření v laboratoři



Obr. 2.10: Grafy průměrné odchylky, měření v laboratoři



Obr. 2.11: Grafy průměrné odchylky, měření v laboratoři

Tab. 2.4: Výsledky měření v reálném prostředí, přehled jednotlivých metod.

Výsledky mikrofonních polí v reálném prostředí						
Generátor rozbití skla				Siréna		
Křížové pole				Křížové pole		
Pohyb ve °	Metoda GCC	Metoda PHAT	Metoda ML	Metoda GCC	Metoda PHAT	Metoda ML
10	13,36	12,62	12,62	44,86	10,94	10,94
20	26,62	26,46	26,46	28,55	17,66	17,66
30	36,41	35,67	35,67	29,61	25,25	25,25
40	39,07	44,29	44,29	-	-	-
50	49,84	46,01	46,01	-	-	-
60	50,97	52,25	52,25	-	-	-
70	59,50	60,42	60,42	-	-	-
80	70,54	72,42	72,42	-	-	-

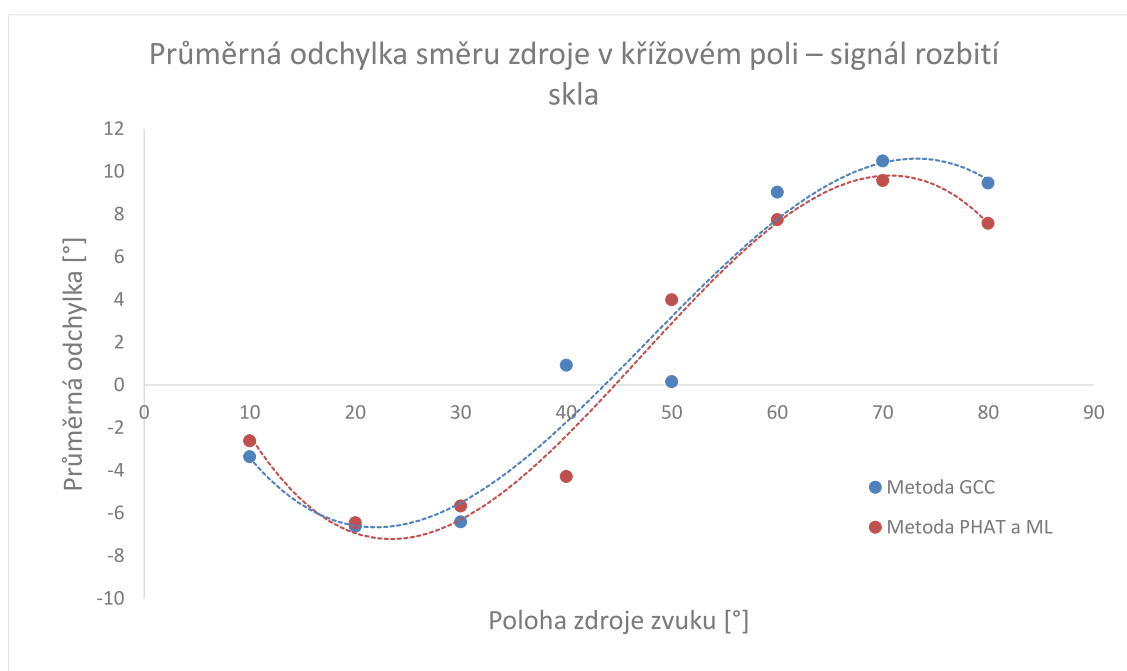
  

Výstřel			
Pohyb ve °	Metoda GCC	Metoda PHAT	Metoda ML
20	24,84	24,69	24,69
40	0,00	0,00	0,00
60	55,52	60,58	60,58

Tab. 2.5: Přehled jednotlivých mikrofónů v poli a jejich přesnost. Metoda GCC, kruhové pole, signál bouchnutí dveří, měřený úhel byl 24°. Červený podklad u čísla značí nejvyšší hodnotu naměřeného úhlu, oranžový potom nejmenší. Průměr v modrém poli byl počítán z absolutní hodnoty úhlů.

NaN	22,01	21,27	15,69	11,74	-26,98	-22,03	-23,40	Pohyb ve °
-22,01	NaN	20,98	0,00	-17,33	-23,45	-22,62	-23,52	24,00
-21,27	-20,98	NaN	-12,05	-22,03	-23,19	-22,83	-23,30	
-15,69	0,00	12,05	NaN	-24,66	-20,78	-23,30	-23,34	
-11,74	17,33	22,03	24,66	NaN	-22,01	-22,34	-21,13	
19,20	23,45	23,19	20,78	22,01	NaN	-18,41	-6,26	33,08
22,03	22,62	22,83	23,30	22,34	33,08	NaN	12,05	-26,98
23,40	23,52	23,30	23,34	21,13	15,91	-12,05	NaN	20,63





Obr. 2.12: Grafy průměrné odchylky, měření v reálném prostředí

### 3 ZÁVĚR

Seznámil jsem se s metodami určování směru přicházejícího zvuku, dále s křížovou korelační funkcí a jejími metodami. V této práci jsem shrnul faktory, které nejvíc ovlivňují výsledek měření. Jedná se o věrnost snímání, dodržení Nyquistova teoremu, filtraci signálu, odstup signálu od šumu, správné zvolení rychlosti zvuku atd. Zjistil jsem vliv vzorkovacího kmitočtu a vzdálenosti mikrofónů na přesnost měření, a to takový, že přímou úměrnost tj. zvětšíme-li dvakrát vzdálenost nebo vzorkovací kmitočet, potom i výsledná přesnost bude dvakrát větší. Dále jsem vytvořil naprogramované testovací signály, kterými jsem simuloval reálné situace, přičemž v laboratoři se tyto výsledky potvrdily (přesnost na jednotlivých mikrofónech při velkém skoku úhlu klesá). Nahrané signály jsem zpracoval a porovnal jednotlivé metody, mikrofonní pole. V bezodrazové komoře nejlépe vycházelo mikrofonní pole křížové. Co se týče metod, PHAT a ML vycházely stejně, což může být chybou úzkého spektra zvuku a velice dobrým SNR, nebo to může být dáno chybou ve zdrojovém kódu, kterou jsem nenalezl. Měření v reálném prostředí proběhlo s šesti mikrofóny, což jistě může znamenat pokles přesnosti, avšak výsledky ukazují lepší výsledky než v laboratoři, což může být dáno nepřesným odhadem správného úhlu nebo průměrováním. Bohužel se mi nepovedlo zprovoznit energetickou metodu, kde nefunguje vybírání úhlů z matice.

## LITERATURA

- [1] SCHIMMEL, Jiří. *Elektroakustika* [online]. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, Purkyňova 118, 612 00, Brno, 2016 [cit. 7. 12. 2016]. ISBN 978-80-214-4716-5. Dostupné z: <[https://www.vutbr.cz/www\\_base/priloha.php?dpid=71624](https://www.vutbr.cz/www_base/priloha.php?dpid=71624)>.
- [2] KHADDOUR, Hassan. *A Comparison of Algorithms of Sound Source Localization Based on Time Delay Estimation*. Elektorevue [online]. Department of Telecommunications FEEC, Brno University of Technology, Purkyňova 118, 61200, Brno, 2011, 1–7 [cit. 2. 12. 2016]. ISSN 1213–1539. Dostupné z:<<http://elektorevue.cz/en/articles/analogue-technics/90a-comparison-of-algorithms-of-sound-source-localization-based\discretionary{-}{-}{-}on-time-delay-estimation-1-1-1/>>.
- [3] CARTER, G. C. *Tutorial Overview of Coherence and Time Delay Estimation*, in *Coherence and Time Delay Estimation—An Applied Tutorial for Research, Development, Test, and Evaluation Engineers*, vol. 1, 1993, pp. 1–27
- [4] KNAPP C. and G. CARTER. *The generalized correlation method for estimation of time delay*, in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, Aug 1976. doi: 10.1109/TASSP.1976.1162830
- [5] VLACH, J., Z. PRŮŠA a P. RAJMIC. *Interaktivní applety pro vybrané operace se signály*. Elektorevue. Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, Purkynova 118, 612 00 Brno, Česká republika, 2009, , 1–13. ISSN 1213–1539. Dostupné také z:<<http://elektorevue.cz/cz/clanky/zpracovani-signalu/0/interaktivni-applety-pro-vybrane-operace-se-signaly-1/>>.
- [6] HLEDÍK, Stanislav. *Fourierova transformace* [online].[cit. 04. 12. 2016]. Ústav fyziky Filozoficko-přírodovědecké fakulty Slezské univerzity v Opavě. Dostupné z:<[http://nora.fpf.slu.cz/~hledik/pub/teach/classes/latex/output\\_samples/BryjovaI-AppsDEFG.pdf](http://nora.fpf.slu.cz/~hledik/pub/teach/classes/latex/output_samples/BryjovaI-AppsDEFG.pdf)>.
- [7] SANTAMARÍA, Ignacio a Javier VÍA. *Estimation of the magnitude squared coherence spectrum based on reduced-rank canonical coordinates* [online]. University of Cantabria, Santander, 39005, Spain, ©2007 [cit. 04. 12. 2016]. Dostupné z:<<https://pdfs.semanticscholar.org/379b/096a034a0c4a87a78dc10c9baa4668cd2ff8.pdf>>.

- [8] HECKBERT, Paul. *Computer Graphics 2: Fourier Transforms and the Fast Fourier Transform (FFT) Algorithm* [online]., 1–13 [cit. 5.12.2016]. Dostupné z:<<https://www.cs.cmu.edu/afs/andrew/scs/cs/15-463/2001/pub/www/notes/fourier/fourier.pdf>>.
- [9] KHADDOUR, Hassan a Jiří Schimmel. *Multiple Sound Sources Localization Using Energetic Analysis Method*. Elektrorevue. [online]. Department of Telecommunications FEEC, Brno University of Technology, Purkyňova 118, 61200, Brno, 2012, 1–9 [cit. 25.05.2017]. ISSN 1213–1539. Dostupné z:<<http://elektrorevue.cz/en/articles/analogue-technics/0/multiple-sound-sources-localization-using-energetic-analysis-method/>>.
- [10] NURUZZAMAN, A., O. BOYRAZ a B. JALALI. *Time-Stretched Short-Time Fourier Transform*. *IEEE Transactions on Instrumentation and Measurement* 2006, 55(2), 598-602. DOI: 10.1109/TIM.2006.864246. ISSN 0018-9456. [cit. 25.05.2017] Dostupné také z: <<http://www.photonics.ucla.edu/media/publication/journal/101109TIM2006864246.pdf>>
- [11] HEINZEL, G., A. RÜDIGER a R. SCHILLING. *Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), including a comprehensive list of window functions and some new at-top windows* [online]. , 1–84 [cit. 25.05.2017]. Dostupné z: <<http://hdl.handle.net/11858/00-001M-0000-0013-5579-7>>
- [12] *Short-time Fourier transform* In: Wikipedia: the free encyclopedia [online], San Francisco (CA): Wikimedia Foundation, 2001- [cit. 30.05.2017]. Dostupné z: <[https://en.wikipedia.org/wiki/Short-time\\_Fourier\\_transform](https://en.wikipedia.org/wiki/Short-time_Fourier_transform)>

# SEZNAM PŘÍLOH

A	Zdrojové výpisy z Matlabu	49
B	Obsah přiloženého CD	52

## A ZDROJOVÉ VÝPISY Z MATLABU

Výpis A.1: GCC metoda naprogramovaná v Matlabu.

```
1 function posun0 = faze(x,x1,d,fs)
2     %funkce se vstupními par.
3     %x je 1. signal
4     %x1 je 2. signal
5     %d je vzdalenost v metrech
6     %fs je vzork. kmitocet
7 c = 340; %rychlost zvuku
8 konst = c/d;
9 if(length(x1)>length(x)) %podmínka pro délku matice
10     x1 = x1(1:length(x));
11 else
12     x = x(1:length(x1));
13 end;
14 X1 = fft (x);
15 X2 = fft (x1);
16 Sx1 = (X1).*conj(X2);
17 fx = ifft (Sx1); %výsledná korelace v počtu vzorků
18
19 index_maxima = find(fx == max(fx),1) - 1; %hledání max.
20 if index_maxima > length(fx)/2
21     index_maxima = (index_maxima-length(fx));
22 end;
23 posun = index_maxima/fs;
24 theta = acosd(posun*konst);
25 plot(fx);
26 disp(['index_maxima: ' num2str(index_maxima)]);
27 disp(['posun[s]: ' num2str(posun)]);
28 disp(['uhel[ $^{\circ}$ ]: ' num2str(theta)]);
```

Výpis A.2: Část kódu naprogramované metody PHAT v Matlabu.

```
1 %predchozi kod je stejny jako u GCC
2 X1 = fft (x);
3 X2 = fft (x1);
4 Sx1 = (X1).*conj(X2);
5 Sx2 = abs(Sx1);
6 Fx = Sx1./Sx2;
7 fx = ifft (Fx);
8 %zpetna transformace, nasleduje hledani maxima
```

Výpis A.3: Část kódu naprogramované metody ML v Matlabu.

```
1 %predchozi kod je stejny jako u GCC
2 X1 = fft (x);
3 X2 = fft (x1);
4 Gx12 = (X1).*conj(X2); %ruzna vykonova spektra
5 Gx11 = (X1).*conj(X1);
6 Gx22=(X2).*conj(X2);
7 delta2 = ((Gx12).^2)/(Gx11.*Gx22); %vykonove spektrum
8 Fiml= (1./abs(Gx12)).*(abs(delta2)./(1-abs(delta2)));
9 %vysledny filtr
10 fx = ifft (Fiml.*X1.*conj(X2));
11 %zpetna transformace, nasleduje hledani maxima
```

Výpis A.4: Jednoduché naprogramování umělého zvuku v Matlabu.

```
1 sum = randn (1,110); %matice sumu
2 sum1 = sum;
3 nuly = zeros (1,7); %matice s nulami
4 sum = [sum,nuly]; %kvuli stejne delce matice
5 sum2 = [nuly,sum1]; %posunutí signálu o vzorek nuly
```

### Výpis A.5: Energetická metoda naprogramovaná v Matlabu

```

1 function energetic =
2   Energeticka_metoda (x, posun, nfft, okno, fs)
3   %počítá směr přicházejícího zvuku
4   % x - vstupní signál o čtyřech mikrofonech
5   % okno - délka analýzy Hammingovým oknem
6   % posun - posun o kolik se chci posunout v čase
7   % nfft - počet FFT bodů; fs - vzorkovací frekvence, Hz
8   x1=x(1,:); x2=x(2,:); x3=x(3,:); x4=x(4,:);
9   %sloupcový vektor
10  Xfb=x1+x2; %výpočty jednotlivých osmičkových mikrofónů
11  Ylr = x3+x4;
12  Wc = ((Xfb.^2)+(Ylr.^2)).^(1/2);
13  %výpočet mikrofónu ve středu
14  z0 = 413; %impedance vzduchu
15  Xfb = Xfb(:); %udělá z x sloupcový vektor
16  % délka signálu
17  delkas = length(Xfb);
18  % periodické hammingovo okno
19  blackhar = blackmanharris(okno, 'periodic');
20  % stft určení matice a rozdělení
21  pocetrad = ceil((1+nfft)/2);
22  %spočítá řádky pro stft, ceil zaokrouhluje k +nekonečnu
23  pocetsloupcu = 1+fix((delkas-okno)/posun);
24  % spočítá sloupce pro stft, fix zaokrouhluje dolů
25  Xvysl = zeros(pocetrad, pocetsloupcu);
26  % vytvoří prázdnou matici o délce sloupců a výšce řádků
27  indx = 0;
28  % index času (kolikátý vzorek rozkládám)
29  % STFT
30  for i = 1:pocetsloupcu
31      xw = Xfb(indx+1:indx+okno);
32      %signál x od časového indexu 1 po časový index + okno,
33      windowf = xw.*blackhar;
34      % FFT
35      X = fft(windowf, nfft);
36      Xvysl(:, i) = X(1:pocetrad);
37      %uložení do matice výsledků
38      indx = indx + posun; %posun indexu
39  end

```



## B OBSAH PŘÍLOŽENÉHO CD

Na přiloženém médiu naleznete naprogramované skripty v Matlabu R2013a. Dále naleznete nahrávky jednotlivých poloh zdroje, jak je popsáno ve stromě souborů. Ve stromě souborů nejsou vypsány všechny nahrané soubory, princip je stejný.

```
/ ..... kořenový adresář přiloženého CD
├── Matlab_skripty ..... naprogramované skripty v Matlabu
│   ├── Nacteni ..... adresář obsahující skripty k načtení zvukových souborů
│   ├── Metoda GCC ..... adresář se skripty GCC metody
│   ├── Metoda PHAT ..... adresář se skripty PHAT metody
│   ├── Metoda ML ..... adresář se skripty ML metody
│   └── Energetická metoda ..... adresář se skripty energetické metody
├── Nahrávky
│   ├── Nahrávky bez mikrofonního pole ..... nahrávky bez mikr. pole
│   │   ├── 01_01 ..... nahrávka č. 1; 1. pozice reproduktoru
│   │   │   ├── Audio 9-001.wav ..... signál na 1. mikrofonu
│   │   │   ├── Audio 9-002.wav ..... signál na 2. mikrofonu
│   │   │   ├── Audio 9-003.wav ..... signál na 3. mikrofonu
│   │   │   ├── Audio 9-004.wav ..... signál na 4. mikrofonu
│   │   │   ├── Audio 9-005.wav ..... signál na 5. mikrofonu
│   │   │   ├── Audio 9-006.wav ..... signál na 6. mikrofonu
│   │   │   ├── Audio 9-007.wav ..... signál na 7. mikrofonu
│   │   │   └── Audio 9-008.wav ..... signál na 8. mikrofonu
│   │   ├── 01_02 ..... nahrávka č. 1; 2. pozice reproduktoru
│   │   ├── 01_03 ..... nahrávka č. 1; 3. pozice reproduktoru
│   │   ├── 01_04 ..... nahrávka č. 1; 4. pozice reproduktoru
│   │   ├── 01_05 ..... nahrávka č. 1; 5. pozice reproduktoru
│   │   └── 01_06 ..... nahrávka č. 1; 6. pozice reproduktoru
│   ├── Nahrávky s mikrofonním polem ..... nahrávky s mikr. polem
│   │   ├── dvere_kriz_pole ..... adresář s nahrávkami zvuku dveří v křížovém poli
│   │   │   ├── 0_stupnu ..... adresář s nahrávkami na pozici 0 stupňů
│   │   │   │   └── Audio-001.wav ..... nahrávka z mikrofonu na pozici 1
│   │   └── 8_stupnu ..... adresář s nahrávkami na pozici 8 stupňů
├── Návrh pole ..... adresář s návrhy polí
├── Zpracování výsledků EXCEL
└── pdf ..... jednotlivé grafické přílohy ve formátu pdf
```

## SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

PHAT	transformace fáze – Phase Transform
ML	pravděpodobnostní maximum – Maximum Likelyhood
GCC	obecná křížová korelace – Generalized Cross Correlation
FFT	rychlá Fourierova transformace – Fast Fourier Transform
DFT	diskrétní Fourierova transformace – Discrete Fourier Transform
STFT	krátkodobá Fourierova transformace – Short Time Fourier Transform
iFFT	zpětná rychlá Fourierova transformace – Inverse Fast Fourier Transform
SNR	odstup signálu od šumu – Signal-to-Noise Ratio
FEKT	Fakulta elektrotechniky a komunikačních technologií
$s(t)$	signál v čase zvukového zdroje
$\psi$	váhovací filr
$w_j$	okno pro krátkodobou Fourierovu transformaci
$f_s$	vzorkovací kmitočet
$p_{00}$	barometrický tlak
$z_a$	jednotková akustická impedance prostředí
$t$	čas
$f$	frekvence
$d$	vzdálenost
$c$	rychlost zvuku